

26 March 1981

# Reagan's non-existent science advisors

Why is the Reagan Administration dragging its feet on the appointment of a Science Advisor in the White House? More than two months have now passed since Dr Frank Press packed his bags and left. From time to time, this or that distinguished person has been mentioned as a possible successor, only to sink beneath the waves of gossip again. The President's budget request dutifully includes provision for a Science Advisor and a small staff, but at the present rate the new fiscal year will have begun before the Administration will be writing cheques against this line item. It is, however, inconceivable that the motive for the delay in appointing a successor to Dr Press is simply parsimony. Even the new stringent budget includes many more extravagances than this modest office. Nor is it likely that the failure to appoint a Science Advisor stems from a lack of qualified candidates. Although some of those mentioned in connection with this office are probably too accustomed to speaking out in public to suit the White House, there can be no enduring shortage of people able to take on the job. Moreover, it is unlikely that President Reagan's White House is fearful that anyone appointed to the post would have to walk in Dr Press's shadow for too long. By common consent, Dr Press did the job superbly, but the new Administration seems not to have been shy of appointing less well known people to other posts previously occupied by relative giants.

The most charitable explanation is that the new Administration has sought to avoid the encumbrance of a Science Advisor while putting together its budget request. And it is easy enough to understand what a nuisance such a person would have been in the past few weeks. The proposal to cancel the United States half of the joint mission, with the European Space Agency, to put a pair of satellites in polar orbits about the Sun would have evoked a string of memoranda explaining that such a course of action might permanently damage the chances of international collaboration on technical projects. Similarly, the Administration may have calculated, the plan to take the National Science Foundation out of science education would certainly have produced a reflex reaction from any Science Advisor anxious to keep his links intact with former colleagues in the wider world. More memoranda would have filtered over from the Old Executive Office. President Reagan's spirits must have sunk quite low at the prospect that a newly appointed Science Advisor, anxious to make his mark alongside the Killians and Wiesners, would have chosen to offer unsolicited and largely unusable (let alone unwelcome) advice about the new Administration's armoury and its lack of policy (so far) on arms control. Why not wait until the summer, when some of the dust will have settled, before inviting such unnecessary problems? On this charitable calculation, the next President's Science Advisor will be tapped on the shoulder some time about midsummer, when his or her plans to spend a vacation at Woods Hole, Aspen or in Europe are too far advanced to be cancelled easily, will take up office in the middle of the next frenzied budget process, and will not start writing awkward memoranda until fiscal 1983

dominates everybody's thoughts. That is the charitable calculation of the new Administration's most charitable motivation.

Unfortunately, more sombre, even sinister calculations are possible. Is there, for example, a chance that the Administration has it in mind never to appoint a successor to Dr Press, not because it seeks to make a modest saving in the budget but because, as a matter of principle, it has decided that it does not want to be encumbered with such a person? Purists will quickly say that no modern administration would harbour such a thought. Has it not been amply demonstrated, ever since the end of the Second World War, that every administration needs a Science Advisor? Did not President Nixon's abolition of the President's Science Advisory Committee (not yet restored, even though Congress would like it back) show what harm can come to those who stab the research establishment in the back? One day, the calculation goes, there will be a constitutional amendment making the office of Science Advisor a part of the machinery of government, and from then on there will be no hesitation. Unfortunately, this simple calculation is wrong. It is probable that President Reagan has persuaded himself that if he cannot find a Science Advisor with all the virtues (distinction, discretion and an unwillingness to write memoranda unless asked), he will do without one. In that, he will be mistaken.

The case for having some kind of Science Advisor is subtle, administrative rather than an issue of principle. It is fashionable to say that science is so important that any president must have his own captive source, but that too stems from a mistaken view of what the world is like. The White House is not short of technical advice. Every government agency in Washington can supply a stream of memoranda to match whatever Dr Press's putative successor writes. The particular function of a Science Advisor in the White House is to hold the ring between the several warring agencies of government that ask contradictory (and impossible) tasks of every government — make unemployment go away, preferably by the encouragement of small business; make universities the powerhouses of innovation but less of a charge on the public purse; ensure peace by some technical device, but provide that if there is trouble in Central Europe, the United States will not be on the losing side. (A Science Advisor worth his salt will also be ready with opinions on whether the MX missile should be based in Utah or at sea, but may not again be required to advise, as in President Kennedy's time, on the best way to grow cranberries in Massachusetts.) In the last resort, however, what matters about the function of the Science Advisor in the White House is not the content nor even the quality of the advice rendered, but the fact of his or her existence. He or she is a lightning-rod, able to concentrate and then harmlessly to discharge the discontents of a powerful constituency. The constituency is not, however, all that powerful, these days, but short of funds and also of friends. An Administration with its wits about it would appoint a Science Advisor now, while it can still choose. If Mr Reagan delays much longer, he will find that he has less choice.

## How not to run telecommunications

The British government seems well on the way to making a hash of its policy on telecommunications. Almost since last July, when the Secretary of State for Industry, Sir Keith Joseph, first spelled out the terms on which the old Post office monopoly would be

relaxed, the House of Commons has been struggling with a nasty compromise of a bill for the reorganization of the British Post Office. The proposal to separate the businesses of handling the mails and of managing telecommunications is sensible, if long



overdue. Most of the other provisions of the bill, especially those concerned with the new telecommunications business (trendily known as British Telecom) are unsatisfactory (see *Nature* 288, 424; 1980). Although British Telecom will no longer (if the bill becomes law) be able to decide for itself how to exercise all of its monopoly powers, these questions will be decided not by the balance of competition in a free market but by the Secretary of State and his civil servants, often with the connivance of British Telecom itself. Thus even on the relatively minor, or at least technical, question of what equipment telephone users may attach to their telephone circuits, the government has now agreed that recommendations by the British Standards Institution should be scrutinized (and no doubt quarrelled with) by British Telecom before being decided by the Department of Industry. That British telephone users will suffer from this fainthearted fudge is clear. Those principally concerned have so far failed to recognize that British Telecom will also be damaged.

Signs of trouble have already surfaced. Last week (19 March), the chairman of British Telecom, Sir George Jefferson, was complaining to an audience of equipment manufacturers that development is being hampered by the lack of capital. The complaint is all too familiar. Even in prosperous times, British governments have been unwilling to let the most buoyant of the nationalized industries modernize its network as quickly as technology would allow. When, as now, the cornerstone of financial policy is to keep a tight rein on public expenditure and thus on the need to borrow from the money markets, and when such funds as can be borrowed are increasingly being used to support industries (private as well as public) in danger of collapse, the pressure on British Telecom is greater than ever. The irony is that the investment even of publicly borrowed funds in the telecommunications network would be profitable, which is more than can be claimed for the other investments of borrowed money which the British government has in the past few months felt bound to make in British Steel, British Shipbuilders and British Leyland. As luck will have it, the continuing restrictions on British Telecom's freedom to borrow capital have now been sharpened because its own revenues have risen less quickly than expected, chiefly as a consequence of the recession. The result will be that British Telecom will have to cut back its already modest plans for development in the next two years. Some years from now, when the recession is eventually ended, Britain will be less well endowed with telecommunications than it could have been. The effect of government policy is thus to starve the future so as to feed the past.

But is not this impediment to the proper development of a communications network merely an unfortunate but unavoidable consequence of the recession? That is the simple response to Sir George Jefferson's plea that British Telecom should be liberated from the shackles that now bind it. The calculation is misleading. Can it ever be wise for a government playing entrepreneur to make unprofitable investments while refusing to invest in profitable projects? For private investors, that would be a path to bankruptcy. For governments, good sense is habitually clouded by political considerations. The British government has been investing in steelmaking, shipbuilding and the manufacture of motor cars in part because it cannot bring itself to face the social upheaval that the collapse of these traditional industries would cause in the parts of the United Kingdom in which they are concentrated. Unemployment, the government seems to have decided, is high enough already. So, too, it is. But to fail to invest in potentially productive enterprises such as telecommunications will entail that, when the recession eventually drags to an end, high unemployment will persist for longer than would otherwise be necessary.

How, in these circumstances, can the British government arrange for the development of the telecommunications network at the pace that technology will allow? For the government, the question does not come out of the blue. Sir Horace Barlow, Sir George Jefferson's predecessor at the Post Office, resigned a year ago because the Department of Industry would not agree to his proposal that the development of the telecommunications

network should be financed by allowing the Post Office to borrow some of its capital requirements from private sources. The Treasury with a simple logic insists that any debt that is ultimately guaranteed by the public purse must be counted as public borrowing. But there are ways round this over-simple rule, some of them already embodied in the Post Office Bill now before the House of Commons. For one thing, to the extent that the shading of the public monopoly will make it possible for the first time (in Britain) for terminal equipment to be purchased by private users and not purchased and leased to users by the monopoly, there is scope for reducing the public cost of telecommunications development. But the bill also includes the delphic provision that British Telecom will be allowed to borrow money from private sources for joint enterprises with private industry in which it does not have a controlling interest. In principle, these arrangements would allow expensive capital projects, the use of earth satellites within the telecommunications network for example, or the development of novel kinds of trunk transmissions routes, to be financed privately and then leased back to British Telecom. Precisely such a device is being considered by the British government as a means of allowing the British railways to be electrified without undermining forecasts of public borrowing in the years ahead. So why, while complaining at the chains that bind it, does not British Telecom point to the ways in which its own development could lighten the government's need to borrow money?

The explanation is straightforward but not creditable. Neither British Telecom nor the Department of Industry appears as yet to have adjusted to the opportunities provided by the shading of the public monopoly on telecommunications. While rightly urging that the pace of development in telecommunications should be faster — much faster — British Telecom seems not to have learned that it cannot do everything itself. If, however, its interest is the rapid development of a modern telecommunications network, it should welcome the chance of having as much as possible of the financial burden carried on other shoulders. It should, for example, welcome the prospect that it may no longer be required to buy terminal equipment from manufacturers simply so as to lease it on to users, tying up capital in the process. In practice, British Telecom has taken a much more defensive line, and is clearly planning to compete as fiercely as it can with such of the suppliers of terminal equipment as the Department of Industry may license in the fullness of time. Moreover, as the bill is now drafted, there is nothing to ensure the new monopoly will compete with upstart suppliers on terms that are objectively fair. Indeed, there is a danger that British Telecom will try to beat off the new suppliers by devoting too much of its own resources to its first essay in competition, thus further hampering its capacity to invest in long-term developments. In the circumstances, British Telecom's pleas for more freedom to borrow money sound more hollow than intended.

The solution must lie with the government, the only participant in the reorganization of the British telecommunications network which has an interest that the old monopoly should be eroded and also that the development of the network should be less dependent on public borrowing. Unfortunately, the Department of Industry cannot ensure that British Telecom behaves sensibly while it has assumed for itself the role of final arbiter on trivial technical questions such as the specification and design of the terminal equipment that may be attached to telephone circuits. What is needed, even at this late stage, is that the British government should step back from the day to day regulation of the telecommunications network, delegating that responsibility to a less partial body able to take a judicial view of what constitutes fair competition between British Telecom and the rest of the telecommunications industry. Such devices have been found necessary elsewhere, in the United States for example, in striking a balance between the public interest and those of public monopolies. For the British government, and for British Telecom, the prize would be some relief from the problem of financing a technological development of great promise in the middle of a recession.



# DNA guidelines for further attenuation

## NIH to consider new voluntary provisions

### Washington

Two prominent members of the National Institutes of Health (NIH) Recombinant DNA Advisory Committee (RAC) are suggesting a major transformation of the guidelines covering recombinant DNA research which would not only significantly reduce containment levels for a variety of experiments, but also convert the guidelines into a non-regulatory "code of standard practice". Dr Alan Campbell of Stanford University and Dr David Baltimore of the Massachusetts Institute of Technology claim that revoking the mandatory aspects of the guidelines is compatible with the experience gained since they were introduced five years ago.

The recommendation from Dr Campbell and Dr Baltimore will be discussed at next month's meeting of RAC. Also under consideration will be some less radical proposals suggested at a recent meeting of the chairpersons of local biosafety committees (*Nature* 288, 529; 1980), focusing primarily on reducing the paperwork involved in registering experiments conducted at the minimal P1 containment level, which is expected to receive the committee's support.

In a letter to Dr Donald Fredrickson, director of NIH, the two authors claim that since 1976, when the guidelines were introduced, neither experimental evidence nor theoretical arguments have been put forward to support the argument that recombinant DNA research poses any danger to human health. According to their proposal, the list of prohibited and exempt experiments would be retained, as would RAC. But all mechanisms which have been set up to enforce the guidelines as regulations — including, presumably, local institutional biosafety committees — would no longer be required.

Dr Campbell and Dr Baltimore argue that singling out some classes of experiments as needing higher levels of physical containment has been made at the cost of discouraging variety and innovation and thereby limiting access to useful knowledge. In contrast, they say, "the benefit is likely to be zero".

If NIH agrees to move in this direction, it is likely to proceed cautiously. At the last RAC meeting in January, the committee approved by nine votes to eight, with three abstentions, a proposal from Dr Winston Brill of the University of Wisconsin that the guidelines be changed to permit

recombinant DNA experiments using nonpathogenic prokaryotes and lower eukaryotes to be conducted at P1 containment.

However, in a notice published recently in the *Federal Register*, Dr Fredrickson says that in the light of the closeness of the vote, he is accepting in preference a compromise proposed by another committee member, Dr Susan Gottesman, which would allow the Office of Recombinant DNA Activities to lower containment levels for such experiments on a case-by-case basis, but without requiring RAC review.

One effect of the proposal of Dr Campbell and Dr Baltimore would be to throw into greater prominence suggested safety guidelines for microbiological and biomedical laboratories prepared by the Center for Disease Control (CDC) — which, like NIH, is part of the Public Health Service. However CDC itself has become a recent focus of controversy, following its publication last autumn of an updated draft of what it calls a "national code of practice".

The document, which has been widely circulated to universities and other research laboratories, is divided into two principal sections; and it is the second

section which has generated the most detailed criticism. This contains proposals for treating several types of organisms in a way which many microbiologists feel would be overly restrictive — in particular using viruses such as SV40. As now written, the guidelines "would put most virologists out of business" says Dr Harlyn Halvorson, chairman of the public and scientific affairs board of the American Society for Microbiology (ASM), which has been coordinating the views of several professional societies and is shortly to issue a detailed critique of the proposals.

CDC officials maintain that they are anxious to achieve a consensus on the guidelines, and plan a thorough revision in the light of the many comments that they receive. They also point out that the guidelines do not have the force of law outside public laboratories, and are intended more as a guide to local biosafety officers.

In reality, however, the CDC guidelines are likely to be widely adopted as a broad framework for dealing with potentially dangerous microorganisms, whether used in recombinant DNA experiments or not. In a letter to CDC director Dr William H. Foege, ASM asks CDC to broaden the basis for discussions with the biomedical research community — something that

## Guidelines transgressor let off lightly

### Washington

Dr Ian Kennedy, discovered last year to have been cloning Semliki Forest virus and not Sindbis virus, has been let off with a caution. The review committee set up by the National Institutes of Health (NIH) to investigate charges that he had infringed regulations covering recombinant DNA research in his laboratory at the University of California in San Diego has decided that no further action can be taken against him. Dr Kennedy is no longer conducting research funded by NIH and has no grant application under review for NIH funding.

The panel has also, however, recommended that, in the light of what it claims was Dr Kennedy's failure to comply with the NIH guidelines covering recombinant DNA research, during the next two years "his use of such techniques may be made subject to such additional conditions as the awarding [NIH bureau, institute or division] may require".

The NIH review committee throws little light on whether, when Dr Kennedy performed experiments using Semliki Forest virus which were at the time proscribed by the NIH rules, he did so intentionally or merely as the result of cross-infection. "It would not be fruitful for NIH staff to engage in an extensive examination of Dr Kennedy's laboratory notebooks or to interview his co-workers in order to attempt a resolution of complicated and disputed chronology."

The committee, nevertheless, faults Dr Kennedy for not seeking prior approval for recombinant DNA experiments in mouse L cells, from which transformed RNA was analysed, and says that Dr Kennedy failed to act responsibly.

Commenting on the report, Dr Donald Fredrickson, director of NIH, said that he accepted the committee's recommendations, and has asked both NIH officials and the university chancellor, Dr Richard Atkinson, to take appropriate action.

The review committee's report comes two weeks after Dr Martin Cline agreed to step down permanently as head of the division of haematology and oncology at the University of California, Los Angeles, following the university's discovery that he had used recombinant DNA molecules in treating patients with thalassaemia without the consent of the university committee responsible for protecting human subjects during clinical experimentation.

A statement issued by the university said that although both the patients and their families were told of the benefits and risks of the experimental treatment — which was carried out on one patient in Italy and another in Israel — neither university nor hospital officials were informed. The university's chancellor, Charles E. Young, added, however, that no harm had come to the patients treated by Dr Cline, whom he hoped would remain at the university.

David Dickson



CDC officials claim they would do.

The question that remains is how this would affect the future of RAC. One option might be to open the committee up into a general "biosafety advisory committee", a proposal already supported by several committee members. Another would be to broaden its scope to consider some of the ethical issues raised by the use of genetic engineering techniques in medical experimentation — a direction which Dr Fredrickson is said to favour. Next month's meeting will provide some clues.

David Dickson

## UK research councils

### Spending too soon

Sir Geoffrey Allen, chairman of the Science Research Council, was let off lightly by the House of Commons Public Accounts Committee last week. The committee had hauled him in to explain why the council had overspent in the year 1979–80 by £4.1 million. Mr Joel Barnett, chairman of the committee, reprimanded Sir Geoffrey for errors of judgement, but seemed to accept his explanation.

The trouble began, said Sir Geoffrey, with the council underspending its budget



No Tower for Allen

by £2 million in each of the previous two years and with the promise in late 1978 from Mrs Shirley Williams, the Secretary of State for Education and Science, of an extra £33 million for the years 1979–80 to 1982–83. To make sure that unspent money would not be taken back, Sir Geoffrey decided to spend the extra allocation for 1979–80 quickly and asked universities to apply for research and equipment grants. But by July 1979 it was beginning to be apparent that the council might be overcommitting funds, especially as the new Conservative government, elected the previous April, had said that it would not honour Mrs Williams's promise in full.

By September things had got worse. The council then knew that it would only receive £5 million of the £7.5 million it had been expecting for 1979–80. An attempt to withdraw some of the extra grants awarded

was unsuccessful because equipment had already been delivered and money spent. By October, the council was forecasting excess expenditure of £7 million by the end of the financial year. In the event, it managed to save £2.9 million, mainly on international subscriptions.

The council had been caught out because equipment manufacturers were able to deliver on time and because universities, with cash flow problems of their own, submitted bills promptly. Work on the spallation neutron source at the Daresbury laboratory had also been completed earlier than expected — the council had been hoping that the bills would fall in the succeeding year.

Mr Barnett suggested that Sir Geoffrey's stratagem to ease cash flow by postponing payments amounted to deception of Parliament. But officials from the Department of Education and Science and the Treasury were more sympathetic in their evidence. The problem is that the council, like other government-funded bodies, has to plan its annual expenditure within a very small margin of error. Sir Geoffrey's request that the error be raised from less than 0.2 per cent to 1 per cent, however, was quashed by the Treasury witness who said that an announcement had been made the previous day that no easing of cost margins was to be granted for any public agency.

The council's overspending for last year is bound to make this year's financial problems even more delicate, especially with the Public Accounts Committee breathing down the council's back. In the meantime, Sir Geoffrey expects to be issued with a formal reprimand. He is unlikely to be sent to the Tower of London.

Judy Redfearn

## Polish academics

### Degrees in question

Edward Gierek, former First Secretary of the Polish United Workers' Party, made a comeback to public life recently — fighting to defend his right to an MSc. In an interview published in the Katowice Party daily, *Trybuna Robotnicza*, Mr Gierek complained that he had been subject to a prolonged campaign of denigration.

Mr Gierek was referring to allegations made by the Krakow Party daily *Gazeta Krakowska*, that the Krakow Akademia Gorniczo-Hutnicza (Mining and Metallurgy Academy — AGH) had improperly awarded scientific degrees to prominent political figures including Zdzislaw Grudzien (former Katowice Party chief), Franciszek Szlachcic and Stanislaw Kowalczyk (former Ministers of the Interior) and to Mr Gierek himself.

Meanwhile, the Solidarity chapter at AGH had published a list of people alleged to have gained degrees improperly, claiming that in the early post-war period special crammer courses were introduced in a number of technical specialities — but

that in practice only Party and government functionaries were able to take advantage of them.

Mr Gierek says that he was only one of several thousand people who had taken advantage of the courses intended to allow people with long industrial experience to get a degree in engineering "without regular studies".

The rector of AGH is Dr Roman Ney, who last autumn became a candidate member of the Party politburo. He was swift to respond to the suggestion of irregularities, and set up a special commission to investigate the allegations. In Gierek's case it found no evidence of malpractice. AGH had on file a copy of his diploma as a mining engineer, specializing in the exploitation of coal deposits, and duly approved by the examination board. Regarding the other allegations that in 1957–61 Grudzien, Szlachcic, Kowalczyk and others had been allowed to complete four-year extramural courses in the economics of metallurgy in just over one year, the AGH department of metallurgy simply reported that these particular courses had been established on the orders of the Minister of Education.

Further queries have been raised about the academic status of Andrzej Jaroszewicz (son of the former Prime Minister). On 9 March, Krakow Technical University set up a special commission to investigate the circumstances surrounding his external degree in engineering.

These swift responses to hints of irregularity are in sharp contrast to the attitudes of officialdom which led, in 1976 to the resignation of the then President of the Polish Academy of Sciences, Professor Janusz Groszkowski, from all his public offices. Professor Groszkowski had attempted to intervene in the case of Mrs Aleksandra Hankus, who in 1964 had complained to the Krakow Technical University that a Mr Tibor Petrys had stolen her scientific results. But the Technical University refused to take any action against Mr Petrys, granted him his DSc and later helped him prepare evidence for a libel suit against Mrs Hankus.

After an 11-year struggle — during which she served an 11-month sentence for libel and was sent by the judiciary for psychiatric tests — the Supreme Court finally vindicated Mrs Hankus, stating that she had proved that Petrys had "usurped her scientific work and cheated in obtaining the degree of Doctor of Sciences". When Professor Groszkowski tried to bring a legal action against Mr Petrys, the Procurator General blocked the case for a year, since, according to Professor Groszkowski, Mr Petrys still enjoyed the support of "some organs of justice". But eventually Mr Petrys' degree was revoked.

In September 1980 Professor Groszkowski at last re-entered public life with an open letter to the academy, calling for a return to true academic standards in Polish science.

Vera Rich

## US universities

**Commercial interest***Washington*

Legislation passed last year by the US Congress to encourage the scientific community to take a greater interest in the commercial exploitation of their research — for example by giving universities and small businesses the right to patents resulting from federally-sponsored research — is already bearing fruit. Spurred by the new patent laws, Stanford University is considering whether to invest university funds in companies set up by faculty members to exploit their research. And the National Institutes of Health (NIH) are expected to reverse a 20-year-old policy which excludes scientists who work for profit-making companies from applying for competitive research grants.

At the same time, there are renewed pressures in Congress for legislation to offer private corporations tax incentives for supporting basic research in universities. One attraction for the universities stems from the increasing constraints on federal funding, but industrial support for university research — at present only about 3 per cent of the total — is unlikely to become a substitute for the federal government.

**New NASA man***Washington*

The Reagan Administration is expected to announce shortly the appointment of Dr James Beggs, at present executive vice-president for aerospace at General Dynamics Corporation, to be the next administrator of the National Aeronautics and Space Administration (NASA). Dr Beggs was NASA's associate administrator for advanced research and technology in the late 1960's and before joining General Dynamics had acted as Under-Secretary in the Department of Energy. His name was informally mentioned by Vice-President George Bush during a visit to the Kennedy Space Centre last week, and a formal announcement about Dr Beggs' nomination is expected from the White House this week.

Meanwhile there is still no news about the post of science adviser and director of the Office of Science and Technology Policy (OSTP). A decision by the White House to include funds for OSTP in its budget request for 1982 indicates that no official decision has yet been made to eliminate the office or its activities. However, there are several rumours in Washington that the science adviser's post may be downgraded, and that OSTP could be absorbed into the Office of Management and Budget, with which it worked increasingly closely during the Carter Administration. **David Dickson**

The Stanford proposal was put to a meeting of the university's committee on research two weeks ago by Dr William Massey, professor of business administration and vice-president for business and finance. In particular, he spoke in favour of setting up a Stanford investment trust which would operate partly on capital provided by the university and partly on capital raised outside. Such a trust would make "modest" equity investments in the early financing of high-technology companies whose products derived in some way from work at Stanford.

Several scientists at the Stanford meeting, among them Dr Edward Garvin, an adjunct professor at the Stanford Linear Accelerator Center, were concerned about the potential impact on research activity of the new proposals, particularly if it meant that the potential ability to earn money for the university became a major factor in making research choices or faculty appointments.

At NIH, the change in rules covering applications for research grants is expected to be announced in the *Federal Register* within the next few weeks. It reverses a restriction applied to all funding by the Public Health Service in 1962, partly as an attempt to ensure that research funds were used to help build up university and medical schools, partly because of some well-publicized abuses.

A return to the situation of before 1962, which would not require ratification by Congress since the decision was essentially administrative, has been under discussion for several years. The document setting out the change in procedures is now said to be awaiting merely the signature of Health Secretary Mr Richard Schweiker, who is not expected to raise significant objections.

The legislation had some important administrative implications. For example, changes in the patent law now required NIH to keep track of all patents awarded to universities from federally-sponsored research, in order that the institutions should be able to exercise "march-in rights" if they consider that a patent is not being sufficiently exploited.

Dr Charles Lowe, head of the institutes' Office of Medical Applications of Research, warned that such changes could have important implications, and steps might be necessary, for example, to protect the integrity of the peer review process. His warnings were echoed by Dr Philip Siekevitz, professor of cell biology at Rockefeller University in New York, who said there was a danger that biomedical research priorities would be shifted from social need to return on investment.

Dr Donald Fredrickson, director of NIH, said he expected that many of the changes described by Dr Lowe would come about and that NIH would in particular be looking closely at the experiences of the National Science Foundation in forging ties between industry and universities.

**David Dickson**

## UK astronomy

**Change of direction**

Responsibility for the construction of the La Palma observatory in the Canary Islands, Britain's hope for good optical astronomy in the Northern Hemisphere, will be handed over to Professor Alexander Boksenberg, now professor of physics at University College London, on 1 October. Professor Boksenberg will then succeed Professor F. Graham Smith as director of the Royal Greenwich Observatory, which is responsible for building and subsequently operating the La Palma site. Professor Smith, a radioastronomer, will succeed Sir Bernard Lovell as professor of radio astronomy at the University of Manchester and director of the Jodrell Bank radiotelescope.

The La Palma observatory, which will house four major telescopes, is likely to be the Greenwich observatory's largest project for a long time to come. The Science Research Council, which made the new appointment, hopes that one of Professor Boksenberg's first tasks will be to chivy the Spanish authorities, who have been responsible for some delay, into speedier action. But Professor Boksenberg sees his main task as encouraging astronomers in British universities to take the fullest possible part in the new observatory. He would like to see their involvement extend beyond the use of the telescopes for observation to the design of instruments and software, either in their own laboratories or at the Greenwich Herstmonceux site.

Since the Royal Greenwich Observatory took on the construction of the La Palma observatory, its own research effort has gradually diminished, a trend that Professor Boksenberg would like to see reversed. He hopes to encourage more astronomers to join his organization and, by collaboration with universities and Britain's other major observatory, the Royal Observatory Edinburgh, to transform the astronomy community into "a happy family".

Professor Boksenberg's previous experience seems particularly relevant to an organization involved in both instrument-building and astronomical research. His expertise in building image intensifying instruments has won him collaboration on several major worldwide telescopes. The image photon counting system, his big success, earned him observation time on the Kitt Peak and Palomar Mountain Observatories in the United States. He is at present building instruments for the faint object camera, Europe's contribution to the space telescope.

His observations have been mainly extragalactic, of quasars in particular. As well as using groundbased optical telescopes, he has also worked with data from and designed instruments for the



International Ultraviolet Explorer satellite. Professor Boksenberg hopes to retain an interest in astronomy at University College London by becoming a visiting professor there. **Judy Redfearn**

## Oak tree wilt

### Europe in dread

*Brussels*

The European Commission is about to launch a research programme to find a cure for the tree disease oak wilt. And a system of import controls is already in force to ensure that wood from the United States does not trigger off an epidemic as devastating as Dutch elm disease.

Experts on both sides of the Atlantic fear that if oak wilt were to reach Europe, it would in the European climate be the death knell of the oak tree. The fungus has spread slowly in the United States and is concentrated in the areas below the Great Lakes. Oak wilt is a fungus (*Ceratostoma fagacearum*) whose spores are thought to be transmitted either by the scullitus beetle or by root grafting (roots intertwining underground). The fungus clogs the tree's vessels, blocking the movement of the sap.

Controversy has surrounded the necessity for and nature of the import regime. Britain and Ireland have been in favour of tight restrictions on imported wood. But aggressive lobbying by the German furniture industry has led to relaxation for some continental industries. The German industry requires oak logs with the bark intact for veneer furniture, and the United Kingdom's Forestry Commission has stressed that the bark poses the greatest dangers of transmission.

As things are, the logs have to arrive accompanied by a certificate to show that the wood comes from uncontaminated areas of the United States. The logs can then enter the Community only through a limited number of ports and cannot then be re-exported to another member state.

Further precautions are nevertheless being taken. The Community's import regime requires that the wood is square-edged, has been dried to remove 20 per cent of the moisture and is subject to heat and air treatment. In case of doubt, the wood can also be ponded or submerged in water to kill remaining fungus. The problem for the furniture industry is that such processes do the wood no good.

The research programme is being mounted in collaboration with American and Canadian research institutes (although Canada is still unaffected). Five contracts are about to be let by the Commission to plant research institutes in the United Kingdom, Germany and Belgium. Details of the programme have not been released, but one contract will go to the Forestry Commission. The total cost of the programme is estimated at 570,000 European Units of Account (£0.54 = 1 EUA) with Europe contributing just over a half, the United States 200,000 EUA and

Canada 50,000 EUA.

Although oak wilt has been known for the past 40 years, work is still going on to map the afflicted areas. Efforts will also be devoted to gaining a greater understanding of how the disease spreads; present evidence suggests that oak wilt would spread far more quickly in Europe than in North America.

Methods exist to kill the fungus in the wood but the more important challenge is to find a way of killing the fungus in the living tree.

Problems have arisen in coordinating the timing of the research, as the Americans are in less of a hurry. European research is also handicapped by the lack of infected wood, for obvious reasons. Relations between Europe and the United States are also coloured by the knowledge that Europeans were responsible for exporting Dutch elm disease to North America in the first place.

**Jasper Becker**

## Soviet higher education

### Candidates change

The Soviet system of higher academic degrees is due for radical overhaul, according to an article in the prestigious *Literaturnaya Gazeta* last week. On the eve of the Annual General Meeting of the Academy of Sciences of the USSR, it published a major critique of the system under which the degree of Doctor of Science is now normally awarded to mature and well-known scientists approaching their fiftieth birthdays.

The author of the article, L. Orlov, is a typical product of the system. He heads a research laboratory employing 50 scientists, and has published more than 80 papers. As a senior scientific worker, holding the degree of Candidate of Science, he earns 300 rubles per month. To obtain the extra 100 rubles a month which go with the title of Doctor of Science, he is now having to rework his research results into a formal thesis, to be defended in exactly the same manner as he defended his Candidate's thesis years ago.

Orlov does not deny that the two-stage degree system had its virtues when it was originally introduced 50 years ago, when the rapid expansion of Soviet science and the lack of trained personnel meant that young scientists who had not yet completed their doctoral research were employed in posts which would normally have gone to Doctors of Science. Under such circumstances, some pre-screening of these "doctoral candidates" was necessary.

Orlov considers the double-examination system no longer relevant since Soviet science and higher education has now reached an extremely high standard. He suggests that the two-stage system is actively harmful to science. Not only do the examiners sometimes let through a young scientist whose thesis is not entirely satisfactory, "since it is, after all, only a

Candidate's thesis", but the need to produce two such pieces of original research in one academic lifetime can seriously hamper a scientist's creativity.

The current Five Year Plan, approved by the Party Congress last month, specifically calls for flexibility in research and development in order to meet the demands of the scientific and technological revolution, and for improvements in training of scientists and workers in higher education. A provocative article in a leading journal is a traditional Soviet way of introducing a possible reform to public opinion, and is usually followed by a spirited "press debate". Orlov's proposals will probably attract considerable sympathy among scientists and academics. Not only would the adoption of a single higher degree eliminate the need to write a second thesis, it would also eliminate the chagrin arising from the present system, which as Orlov notes, gives formal parity between a Soviet "Senior Scientific Worker" who at forty plus has gained his doctorate after 15 years hard work, and some "27-30-year-old young Doctor of Science from East or West Germany".

**Vera Rich**

## Irish research and development

### Marching to Georgia

Ireland is to set up a new contract research institution in the hope of persuading the 800-odd foreign companies based — tax-free — in the country to do some research and development there at last. The new institution, announced last week, and to be called the European Research Institute of Ireland (ERII), will be set up in Limerick with the help of America's Georgia Institute of Technology and will offer 50 per cent government grants to companies wishing to place contracts with it.

Ireland's Industrial Development Authority, which is responsible for the arrangements, foresees ERII as leading Ireland into its third stage of industrial development. The first was purely domestic; the second involved the establishment of foreign industry; and the third will involve research and development placing Ireland among the modern technical countries. The problems in reaching that third stage have been first, that most Irish researchers travel abroad to do their work, to America or Europe, and second, that whereas other countries offered tax incentives to companies to invest in research and development, Ireland could offer none as the firms paid no tax anyway.

The new institute is seen as the solution to these problems, with the grants substituting for tax relief, and high salaries to attract expatriate researchers back to Ireland. The Industrial Development Authority calculates that it will be paying staff at Limerick enough to match the



standard of living they might enjoy if they were based, say, at the Battelle Institute in Frankfurt.

ERII will start formally on 1 June, when two members of the Georgia Institute of Technology Engineering Experiment Station begin work in Limerick. Initially they will be housed in Ireland's National Institute for Higher Education, where president Dr Edward Walsh is said to be playing an "aggressive" part in forging links between university and industry. Recruitment is under way for eight engineers, scientists and computer specialists, but in three years' time, when ERII will move to its own buildings, it is planned that staffing levels should increase to 80.

To finance ERII, the Industrial Development Authority is paying for the buildings (£1.5 million); and the balance will come from such sponsors as Guinness Ireland, the Bank of Ireland, Allied Irish Banks and the industrial Rohan Group. Nevertheless the authority is seeking another £500,000 to launch the institute "effectively", and would like more companies to contribute to the consortium.

**Robert Walgate**

## UK universities

### Jobs to be cut?

One week after the British government's budget, the University Grants Committee has plucked up courage to make some guesses on the quantitative implications for universities. Dr Edward Parkes, the chairman, told the Public Accounts Committee of the House of Commons last week that to meet an 11 per cent cut in real terms in the universities' grant over the next three years, 3,000 academics and 4,000 other staff would have to lose their jobs. Redundancy payments to academics, said Dr Parkes, could be somewhere between £40,000 and £80,000, although realistic figures would not be known until cases had been through the courts. The total cost to the universities could be £100–200 million.

Dr Parkes has since said that in addition to 3,000 redundancies, 2,000 jobs would have to be lost through natural wastage. He estimates that 3,000 academics will leave their posts in the next three years, either through retirement or to take posts else-

where. His current plan allows for 1,000 vital posts to be refilled. A total loss of about 5,000 academic jobs represents just under 15 per cent of the academic workforce. The percentage of job losses will have to be greater than the percentage cut in the budget because economies in other areas of expenditure are virtually impossible, according to Dr Parkes.

The Association of University Teachers disagrees with that view. It is preparing a scheme to avoid large numbers of redundancies by relying on natural wastage, an early retirement option at 50 and spending one-third of the equipment grant on staff. But the scheme is unlikely to appeal to the grants committee which holds that freezing all posts would be the worst option. The problem is that those departments with the highest rate of turnover of staff — medicine, science and engineering — would suffer most, and in many cases, those are the very departments that the government would like to see least damaged. The committee prefers to maintain minimum flexibility in making new appointments and create the minimum number of redundancies.

The Science Research Council, which provides research grants to individual academics, is particularly concerned that good research should not disappear. Ever optimistic, it says that it may be able to support a few outstanding individuals who would otherwise have to go. Such support, however, is unlikely to be very significant.

The Committee of Vice-Chancellors and Principals has estimated that the grant for 1983–84 will be £80 million less than that for this year and that the loss from overseas students' fees could remove a further £70 million, a total sum short of conservative estimates of redundancy payments. The estimated costs may well have come as a surprise to the government, which is showing signs of reconsidering its plans. Sir James Hamilton, Permanent Secretary at the Department of Education and Science, told the Public Accounts Committee that his department was still attempting to work out with the University Grants Committee how the economies could be achieved. Although the budgets for 1981–82 and 1982–83 are fixed, he said, the figure quoted for 1983–84 is still only an estimate, suggesting that there may be room for further negotiation.

Universities, however, are unlikely to make up lost income by taking in more students. Contrary to the impression given in the white paper on public expenditure and by Mr Mark Carlisle, Secretary of State for Education, Sir James Hamilton told the Public Accounts Committee that the government will not increase the grant to universities that overstep their quota of students. The plan is that student numbers should remain roughly constant, implying greater competition for places as the number of eighteen-year-olds in the population continues to increase over the next few years.

**Judy Redfearn**

## Bulgarian astronomers look up and abroad

Bulgaria's new National Astronomical Observatory on Mount Rozhen, opened on 13 March 1981 by President Todor Zhivkov, was several times referred to in the inaugural speeches as a "pre-Congress gift to Bulgarian science". The term is appropriate, for the theses for the forthcoming Congress of the Bulgarian Communist Party state that in the new Five-Year Plan "primary importance should be paid to fundamental research".

This is in sharp contrast to the policies of many other East European countries, which tend to give first priority to application. According to Kiril Serafimov, head of the Bulgarian Committee for Space Research and Exploitation, "we are also a European country and understand that without fundamental science nothing is possible".

Speaking at the opening ceremony, Dr Serafimov inevitably referred to Bulgaria's "thousand-year-old" interest in astronomy — the country is this year celebrating its 1,300th anniversary of statehood. The observatory has three telescopes and a large data-processing and computer unit. The main telescope has a 200-cm hyperbolic mirror with a Coude-Richey-Cretien focusing system and will be used primarily for work on binaries and variable stars. The second, a 70-cm Schmidt instrument (effective diameter 50 cm) with a 4° x 4° field of view, will be used for preliminary search and scanning work. The smallest, a 60-cm instrument, is especially adapted for photoelectric work and brightness determinations.

The observatory is the most southerly in the socialist bloc (and, according to Bulgarian claims, the second largest optical observatory in Europe). The Rozhen team

expects to cooperate closely with the Soviet observatories, especially the Crimean Astrophysical Laboratory and the Special Astrophysical Observatory in the Caucasus, on approximately the same latitude. Intra-Comecon cooperation is well represented at Rozhen — the research programme has been worked out in collaboration with teams from the Soviet Union, the German Democratic Republic and Poland. (The optical equipment has been supplied by Carl Zeiss (Jena).) There has also been some French participation in drawing up the programme, while the data-processing equipment is largely of Western origin.

**Vera Rich**



*Mount Rozhen in snow*

# CORRESPONDENCE

## Toxicity testing

SIR — The dramatic fall in Fisons' shares as a result of the withdrawal of Proxicromil emphasizes the high risks involved in drug development (*Nature* 12 March, p.81). Drug development programmes depend on animal tests for assessing long and short term toxic effects which are notoriously unreliable and expensive.

The effects of drugs on animals cannot be assumed to be the same in humans, irrespective of whether results indicate that a drug is "safe" or "harmful". Legislators should accept this fact and release toxicologists from the present "check-list" testing system which has replaced scientific initiative with standard tests of dubious relevance which are carried out on ever increasing numbers of animals.

On scientific, economic and humanitarian grounds current toxicity testing protocols should be refined to ensure the most efficient use of animals and to incorporate valid non-fundamental screening tests wherever possible.

LESLEY SARJEANT

*Fund for the Replacement of Animals in Medical Experiments,  
London, UK*

## Inherited tolerance?

SIR — I would like to suggest a more straightforward explanation of the fascinating findings of Gorczynski and Steele<sup>1</sup> than the one they have opted for. There is a strong possibility that mouse semen contains some cells other than sperm — lymphocytes, for example. In the mouse, semen is ejaculated intra-cervically, rather than intra-vaginally as in humans. Thus any non-sperm cells will be directly introduced into the uterus in the mouse. Not much is known about the fate of such cells although it appears that sperm are rapidly removed by phagocytosis. The pregnant uterus, however, is to some extent tolerant of foreign cells. Short and Yoshinaga<sup>2</sup>, Wilson<sup>3</sup> and Lions<sup>4</sup> have described experiments which show that heterologous tumour cells can persist for several days in the rat uterus and that these can enter the mucosa only at the time of implantation in pregnancy. Thus the endometrium becomes receptive to foreign cells at the time of implantation, and so non-sperm cells present in semen could presumably become established in the endometrium and might then enter the embryo after implantation.

If we then suppose that tolerance involves clones of cells in a manner analogous to immunity and that some of the non-sperm cells are competent for the transmission of the tolerant state, the embryo might be rendered tolerant to the same antigens as the father. Moreover, if the chances of such an invasion of the embryo were less than 100 per cent, tolerance of a particular antigen would be found in only a proportion of the embryos and some would be tolerant to both of two antigens whilst others would be tolerant only to one. Since the postulated cells conferring tolerance are differentiated they would play no part in embryogenesis but would be carried and perhaps multiplied in the embryo. In subsequent generations, the chances of transmission of the cells would be finite and

the incidence of tolerance to specific antigens would decline. In the mothers, the return to the non-pregnant state after parturition would, presumably, involve the rejection of any foreign clones of cells in her tissues. Thus, subsequent matings to non-"tolerized" fathers would not result in the transmission of previous tolerant states.

In this way, the essential aspects of the Gorczynski and Steele experiments might be explained along rather less controversial lines than at present. At the very least, the possibilities raised here must be eliminated before we venture to compromise the doctrine of Weismann and begin resurrecting the long dead Lamarck.

*Department of Genetics,  
University of Edinburgh,  
Edinburgh, UK*

K.W. JONES

1. Gorczynski, R.M. & Steele, E.J. *Nature* **289**, 678–681 (1981).
2. Short, R.V. & Yoshinaga, K. *J. Reprod. Fert.* **14**, 287 (1967).
3. Wilson, I.B. *Proc. zool. Soc. Lond.* **141**, 137 (1963).
4. Lions, J. thesis, Univ. Cambridge (1970).
5. I thank Johnathan Aitken for discussion of these ideas and for the references cited.

## Exhibit dismay

SIR — I have been following with great interest and some dismay the extended correspondence in *Nature* about the nature and significance of new exhibits in the British Museum (Natural History). I wish to comment on recent letters from three colleagues and fellow-countrymen of mine<sup>1–3</sup>, one British colleague<sup>4</sup> and one French colleague<sup>5</sup>.

McKenna<sup>1</sup> and Nelson<sup>2</sup> speak out against Halstead and in favour of cladistic or Hennigian systematics, to which a considerable part of the new dinosaur exhibit is devoted. Both McKenna and Nelson are now at the American Museum of Natural History, with which I was associated for many years. They belong to a small group now there which is representative neither of the staff of the museum as a whole nor of the majority of American palaeontologists, who, like me, reject Hennigian procedures.

Gould<sup>3</sup> is at the Museum of Comparative Zoology, with which I was also associated for some time. He proclaims that he is not a cladist and "did not care for" the cladistic exhibits at Cromwell Road. He also claims that his theory of punctuated equilibrium is not a Marxist plot and notes that the co-author of the theory, Eldredge, is not a Marxist. I do not wish to discuss Marxism in this connection, but I do note that the punctuated equilibrium of Eldredge and Gould is only one facet, taken to an extreme, long since incorporated in the synthetic or if you like, neo-Darwinian, theory of evolution.

Howgate<sup>4</sup> named me as being "by no means averse to the idea of leaps," that is, to saltatory evolution. He quotes as evidence a statement by me in a book published in 1949 (not 1976 as he has it) that rates of evolution in the horse family varied considerably. I did not state or imply that the rates were ever saltatory, that is, instantaneous.

Janvier<sup>5</sup>, of the Laboratoire de Paléontologie des Vertébrés of the University of Paris claims that cladism is "directly descended from Darwin's ideas on phylogeny and systematics." In context this implies that Darwin was a saltationist, which is not true.

He also notes that some French scientists (Grassé as an example) are anti-Darwinian. In fact French biologists, with the exception of a few radicals, have always been anti-Darwinian, at times for no clear reason other than chauvinism.

G.G. SIMPSON

*Department of Geosciences,  
University of Arizona, Tucson, USA*

1. McKenna, M.C. *Nature* **289**, 626 (1981).
2. Nelson, G. *Nature* **289**, 627 (1981).
3. Gould, S.J. *Nature* **289**, 742 (1981).
4. Howgate, M. *Nature* **289**, 344 (1981).
5. Janvier, P. *Nature* **289**, 626 (1981).

## Evolving theory

SIR — The letter from H.W. Ball and 21 colleagues (*Nature* 12 March, p.82) contains the following astounding statement: "But the theory of evolution would be abandoned tomorrow if a better theory appeared".

The lapse of time between the appearance of a "better theory" and the abandonment of its predecessor can be established with fair accuracy in many topics of science. The history of science provides a large number of such events. Accordingly, the lapse of time between the appearance of the new and the relinquishing of the old might be regarded as one of the many facts of science. Your 22 correspondents, apparently seriously, estimate the magnitude of this lapse of time as one day. One day! They declare their devotion to fact and their abhorrence of prejudice. Are their other facts of a similar colour I wonder, or does evolution evolve more rapidly than any other field of science?

R.V. HESKETH

*Berkeley Nuclear Laboratories,  
Berkeley, Glos., UK*

SIR — Bertolt Brecht put these words into the mouth of Galileo<sup>1</sup>:

"And what we discover today we shall wipe off the slate tomorrow and only write it up again once we have again discovered it. And whatever we wish to find we shall regard, once found, with particular mistrust. So we shall approach the observation of the sun with an irrevocable determination to establish that the earth does *not* move. Only when we have failed, have been utterly and hopelessly beaten and are licking our wounds in the profoundest depression, shall we start to ask if we weren't right after all, and the earth does go round."

As far as evolution is concerned, I think most biologists would agree that "the earth does go round", there being no better way of accounting for the facts *at present*. So how are we to understand the over-cautious round robin from the British Museum? There is a world of difference between "If the theory of evolution is correct . . .", and ". . . evolution is a fact, proven to the limits of scientific rigour." Most of us, like yourself, would wish to be more positive than the first but less dogmatic than the second. In these days of proliferating crazies, the British Museum should not risk giving them more ammunition.

P.L. MILLER

*Queen's College, Oxford, UK*

1. Brecht, B. *Life of Galileo* (translated by J. Willett) (eds Willett, J. & Manheim, R.) (The Collected Plays Vol. 5, pt 1, Eyre Methuen, London, 1980).



## NEWS AND VIEWS

# Genes pieced together — exons delineate homologous structures of diverged lysozymes

from Peter J. Artymiuk, Colin C.F. Blake and Albrecht E. Sippel

GILBERT<sup>1</sup> has proposed that the arrangement of eukaryotic genes in a mosaic of introns and exons might facilitate the evolution of new proteins. He suggests, for example, that if exons correspond to functions, recombinations within introns could put together new proteins with new combinations of functions. Blake<sup>2</sup> has pointed out that if exons also correspond to folded protein structures the new protein is more likely to achieve structural stability by being a combination of folded sub-structures.

If these ideas are correct, we would expect to discover in different proteins folded units of common structure and function, derived from common exons, combined with units of different structure and function. Examples may already be to hand in the dehydrogenases<sup>3</sup>, and in triose phosphate isomerase, pyruvate kinase<sup>4</sup> and KDPG-aldolase<sup>5</sup>, but in the absence of the details of their exon/intron patterns this cannot be confirmed. The definition of four exonic regions in the hen lysozyme molecule<sup>6</sup> allows a comparison of these regions with the regions of structural and functional homology in T4 phage lysozyme<sup>7,8,10</sup>. This comparison is, intriguingly, consistent with the hypotheses of Blake and Gilbert, although as an isolated first attempt it cannot be adduced as proof.

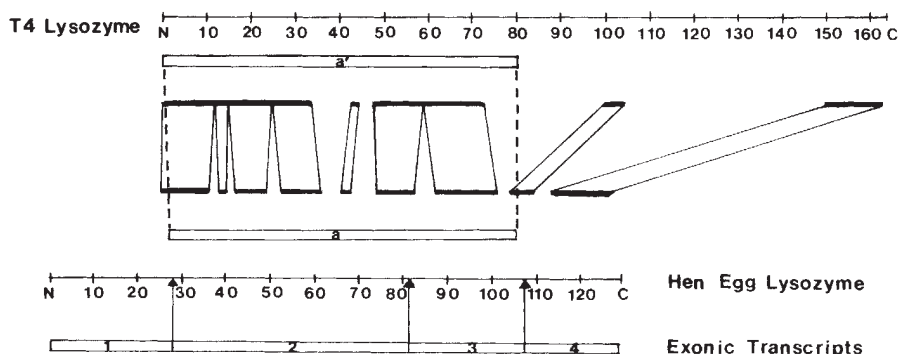
Determination<sup>9</sup> of the structure of T4 phage lysozyme did not at first suggest that there was a close resemblance to hen egg-white lysozyme. Rossmann and Argos<sup>7</sup> later showed that substantial portions of

the molecules were structurally similar. A number of small deletions in one molecule (or insertions in the other) made residues 1–73 of T4 lysozyme overlap with residues 25–101 of hen egg-white lysozyme. Two large insertions in the T4 chain brought out homologies in residues 100–105 and 150–163 of T4 with residues 104–109 and 113–126 of hen. The structural equivalence is shown schematically in the figure. In all, 78  $\alpha$ -carbons were shown to be equivalent with a r.m.s. difference of 4.1 Å. Using somewhat different techniques that do not allow insertions or deletions in the chains, Remington and Matthews<sup>8</sup> confirmed the similarity and showed that residues 27–109 could be made to overlap with residues 1–80 of T4 lysozyme with a r.m.s.  $\alpha$ -carbon difference of 6.0 Å. It is important to note that the structurally equivalent regions correspond to, at most, two-thirds of the hen lysozyme molecule and less than half the T4 lysozyme molecule, and for each lysozyme large contiguous regions of chain have no equivalent in the other molecule.

Rossmann and Argos<sup>7</sup> noted that the equivalent regions encompass the active site regions of each of the molecules and showed that the binding sites for an *N*-acetylglucosamine monomer were only 0.8 Å apart in the two enzymes. Recently Matthews and his colleagues (see this issue of *Nature* p.334) have extended this

positional similarity to a bound tri-*N*-acetylglucosamine moiety, and to Glu 11 and Asp 20 of T4 lysozyme with the catalytic Glu 35 and Asp 52 of hen lysozyme<sup>11</sup>. Furthermore, they suggest that sugar C is bound in a similar mode in the two enzymes and that sugar D is distorted in T4 lysozyme by a close approach of its 6-hydroxyl to protein groups similarly positioned with respect to the binding groups of sugar C, as in hen lysozyme<sup>11</sup>. These observations imply that the structural similarity is accompanied by a similarity in substrate binding and catalytic mechanism. Although each separate piece of evidence is circumstantial, their combination argues strongly for an evolutionary relationship, or more unlikely, an extreme case of convergence in structure and function between the two lysozymes.

Sippel and his colleagues<sup>6</sup> have recently determined the exon pattern in the hen lysozyme gene. They have found that the gene contains four exons, corresponding to residues –18 to 28 (the signal peptide of prelysozyme and residues 1–28 of mature lysozyme), 28–81/82, 81/82–108 and 108–129. They note that although the exon pattern is related to some degree to the structural division of the hen lysozyme molecule, its relation to the functional units of the enzyme is better established. For example, exon 2, which codes for residues 28–81/82, includes the catalytic residues and a cluster of amino acids that bind sugars C, D, E and F of the oligosaccharide substrate and could



A schematic diagram showing the relation between hen egg lysozyme and T4 phage lysozyme. The solid bars in the centre show the residues which are considered equivalent by Rossmann and Argos<sup>7</sup>. The rectangles marked a and a' define the 80-residue segments aligned by Remington and Matthews<sup>8</sup>. The lower rectangles and arrows mark the exonic transcript regions defined by Jung *et al.*<sup>6</sup>.

perhaps function as a primitive glycosidase. Exon 3, which codes for residues 81/82–108, corresponds to a protein fragment that, combined with the exon 2 product, could give additional substrate specificity, determine the cleavage frame for the alternating NAG–NAM chain and increase the catalytic efficiency of the active centre. Exons 1 and 4, on the other hand, code for protein fragments that are not directly involved in the catalytic function, but may increase the stability of the hen lysozyme molecule and/or carry some as yet unknown biological functions of the enzyme.

The regions of structural homology in T4 and hen lysozyme described by Rossmann<sup>7</sup> and Matthews<sup>8</sup> can now be compared with the exon pattern of hen lysozyme defined by Sippel<sup>6</sup> (see the figure). First, it is clear that residues 1 to 24 or 26 (depending on the comparison) of hen lysozyme, corresponding closely to exon 1 (residues 1–28), have no counterpart in the T4 lysozyme molecule; that is, in T4 lysozyme exon 1 is absent. Second, the principal region of overlap between the two lysozymes corresponds closely to the combination of exons 2 and 3 (residues 28–108). In Rossmann's analysis this region corresponds to residues 25–101 of hen lysozyme and residues 1–73 of T4 lysozyme, while Matthews' analysis<sup>8</sup> slightly extends and alters the overlap to include residues 27–109 of hen lysozyme and residues 1–80 of T4 lysozyme. Rossmann's analysis<sup>7</sup> also lines up residues 104–109 of hen lysozyme with residues 100–105 of T4 lysozyme, but this requires a 25-residue insert in the T4 sequence to incorporate it in the major overlap. Accepting some ambiguity with regard to residues 102–109 of hen lysozyme, the correspondence to the combination of exons 2 and 3 is close. If Rossmann's analysis<sup>7</sup> is restricted to these two exons, the overlap reduces the r.m.s.  $\alpha$ -carbon difference from 4.1 Å to 3.2 Å. Finally, correspondence to exon 4 (residues 108–129) of hen lysozyme seems less clear-cut than for exons 1–3. Matthews<sup>8</sup> does not match it with any part of the T4 lysozyme. Rossmann and Argos<sup>7</sup> overlap residues 113–126 of hen lysozyme with residues 150–163 of T4 lysozyme but with a rather high r.m.s.  $\alpha$ -carbon difference of 6.2 Å. In these circumstances it is difficult to tell whether there is or is not a region of

the T4 lysozyme molecule that corresponds to exon 4 of hen lysozyme.

In conclusion, with the exception of exon 4, the agreement between the exons of hen lysozyme and the structural homology with T4 lysozyme is very close. Exon 1 is absent from the T4 molecule and the principal region of overlap corresponds to exons 2 and 3. In view of Sippel's conclusion<sup>6</sup> that exon 2 could represent a primitive glycosidase minigene on which substrate specificity was conferred by the product of exon 3, it is tempting to see exons 2 and 3 as the functional nucleus of the lysozyme molecule. The presence of these two regions in T4 lysozyme could account for the similarity in substrate binding and catalytic mechanism suggested by Matthews<sup>10</sup>. In addition, the absence of exon 1, and the possible absence of exon 4, are not unreasonable if, as Sippel suggests, they play a role in the stability, rather than catalytic function, of the hen lysozyme molecule. Their stabilizing influence may be taken over in the T4 lysozyme molecule

by the large C-terminal lobe or domain, which has no equivalent in the hen molecule. Indeed, Matthews<sup>10</sup> suggests that this C-terminal region of T4 lysozyme has a function in binding the cross-linked peptide of *E. coli* cell walls, which is not necessary for hen lysozyme.

Although this analysis must be regarded as provisional until other proteins can be examined from a similar point of view, this comparison of T4 and hen lysozyme is fully consistent with the view that recombination within introns can rearrange functional exonic regions into new patterns in new protein products. What is most encouraging is the promise that protein structural analysis can help us identify and predict exons in genes and trace evolutionary events at the level of DNA, and conversely that sequence analysis of gene DNA can help us to understand the structure, function and evolution of the final protein product. □

We are grateful to Professor B. W. Matthews for letting us see his manuscript before publication.

## Waves, turbulence, solar convection and ozone

from Michael McIntyre

THE Earth's atmosphere and oceans are increasingly being used as an outdoor laboratory, not only providing information about the behaviour of our environment but also insight into certain difficult, unsolved problems in fluid dynamics. These involve wave-like and turbulent fluid motions at huge Reynolds numbers which are inaccessible to both indoor laboratory experiment and simulation on the world's largest computers.

Empirical information is specially important for the ill understood problem of turbulent motion, and one paper given at a recent conference\* was of unusual interest in this respect. J.J. Finnigan (CSIRO, Canberra) and F. Einaudi (Georgia Institute of Technology, Atlanta) have produced what may be the first really detailed information about how an internal gravity wave interacts with pre-existing turbulence. In the case they analysed the wave was nearly monochromatic, and by good fortune had propagated past the 300 m tower and ground-level microbarograph array at the Boulder Atmospheric Observatory, Colorado at a time when a large volume of data was being taken. Turbulent and wave-induced fluctuations could be separated cleanly, revealing how

the turbulent Reynolds stresses varied with the wave motion. To a first approximation the turbulence behaved quasi-elastically, indicating that rapid-distortion turbulence theory should apply. The observed behaviour is quite different from what would have been expected from 'eddy viscosity' or 'mixing length' ideas. It is encouraging to see further direct evidence for the use of rapid-distortion theory, which is complete and rational within its limits of applicability. Other recent lines of evidence come from the work of J.C.R. Hunt and his collaborators at Cambridge on turbulent flow over hills and around buildings, for which rapid-distortion theory is proving to be more useful than previously supposed, and from an estimate by D.O. Gough (Institute of Astronomy, University of Cambridge) using recent observations of the solar 'five-minute' oscillations, showing clearly that they must interact with the Sun's turbulent convection zone in a similar way.

As suggested some years ago by A.A. Townsend, the same ideas seem likely to be important also for understanding the generation of ocean waves by wind. Other aspects of that celebrated and still unsolved problem were illuminated by E. Mollo-Christensen (MIT) and A. Ramamonjia-

1. Gilbert, W. *Nature, News and Views* 271, 501 (1978).
2. Blake, C. C. F. *Nature, News and Views* 273, 267 (1978).
3. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* 250, 194 (1974).
4. Levine, M., Muirhead, H., Stammers, D. K. & Stuart, D. I. *Nature* 271, 626 (1978).
5. Richardson, J. S. *Biochem. biophys. Res. Commun.* 90, 285 (1979).
6. Jung, A., Sippel, A. E., Grez, M. & Gehütz, G. *Proc. natn. Acad. Sci. U.S.A.* 77, 5759 (1980).
7. Rossmann, M. G. & Argos, P. *J. molec. Biol.* 105, 75–96 (1976).
8. Remington, S. J. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* 75, 2180 (1978).
9. Matthews, B. W. & Remington, S. J. *Proc. natn. Acad. Sci. U.S.A.* 71, 4178 (1974).
10. Matthews, B. W., Grütter, M. G., Anderson, W. F. & Remington, S. J. *Nature* 290, 334 (1981).
11. Blake, C. C. F. *et al. Proc. R. Soc. B167*, 378 (1967).

\*The American Meteorological Society's third conference on 'Atmospheric and Oceanic Waves and Stability' was held at San Diego from 19 to 23 January, concurrently with AMS meetings on the 'Upper Atmosphere', 'Climate Variations' and the 'Economic and Social Value of Weather and Climate Information'.

Michael McIntyre is in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge.



risoa (University of Aix-Marseille), who presented new laboratory evidence of strongly nonlinear coupling between the air flow and modulational instabilities of the waves. More light was thrown on the instabilities themselves by a remarkable series of photographs and measurements by Y.-M. Su and A.W. Green (NORDA, NSTL Station, Mississippi) showing a succession of instabilities in a train of moderately steep, monochromatic waves. The waves were generated by a wavemaker about 16m long in an outdoor basin 100×340m. A small-scale, essentially three-dimensional instability, very regular in form but as yet unexplained theoretically, was prominent. Its nonlinear evolution appeared to culminate in scattering of some of the wave energy in a direction about 30° away from that of the primary wavetrain, a phenomenon hardly likely to have been identified in a narrow laboratory tank.

Most of our knowledge about the large-scale, low-frequency oscillations known as Rossby waves comes from observations of the atmosphere and ocean, where in practice the waves often have large amplitudes and occupy the no-man's land between 'waves' and 'turbulence'. Experimental possibilities are severely restricted not only by small dimensions but also by the constancy of gravity in the laboratory. Rossby waves can be expected to occur also in the stably stratified interiors of rotating stars. Y. Q. Kang and L. Magaard (University of Hawaii) produced what appeared to be the clearest example yet observed of mid-latitude, oceanic Rossby waves, in a 'baroclinic' mode unobtainable in the laboratory. The wavelength was about 300 km and the waves had a phase velocity towards the north-west and group velocity towards the south-west with magnitudes of the order of a centimetre per second. A point of special interest was that the waves had extremely large amplitudes, as measured by the ratio of a typical particle velocity to a typical phase velocity — well in excess of the kind of amplitudes at which most waves in fluids become very strongly dissipative.

Rossby waves in the atmosphere, too, were a topic of renewed interest. A number of papers were on theoretical models related to the recent work of J.G. Charney (MIT) and his collaborators, in which global-scale, topographically generated Rossby waves are allowed to interact with an azimuthal mean flow driven by a constant forcing. This simple interaction seems capable of explaining the observed tendency of the atmospheric circulation to flip between quasi-stable states, with large and small amplitudes in those wave components which are phase locked to the topography. The large-amplitude states are associated with 'blocking' patterns and persistent spells of anomalous weather. In winter the same Rossby waves can penetrate the stratosphere, where under the right conditions they trigger 'sudden

warmings' — sporadic but spectacular upheavals in the state of the polar stratosphere in which dynamically induced subsidence of air over the polar cap causes large, adiabatic temperature rises. Whether or not a sudden warming occurs is sensitively dependent on conditions in the stratosphere itself, as well as in the troposphere. At the concurrent meeting on the 'Upper Atmosphere' there were signs that our understanding of the conditions favouring sudden warmings may be about to improve significantly. Recent progress in the dynamical theory, together with the uniform global coverage offered by infrared satellite observations, have led to better ways of analysing the observations

and more illuminating comparisons with computer simulations. Some real clues are emerging as to how and why the behaviour of the stratosphere differs from that of the computer simulations, and we can await the results of current modelling efforts with considerable interest. The sudden warming is a large-scale dynamical experiment which nature kindly performs from time to time. It is one of the crucial tests of our ability to simulate the large-scale dynamics of the stratosphere. The dynamics, in turn, are one of the necessary ingredients in attempts to understand the stratospheric circulation in general, and to predict the effects of pollution on the ozone layer in particular. □

## Climatic variation in San Diego

from M.J. Salinger and C.B. Sear

DURING anomalously warm and dry weather in San Diego in January, climatologists met to discuss climatic variations. Much of the meeting\* centred on the implications for global climate of the likely increase in atmospheric CO<sub>2</sub> concentration during the next few decades, but other topics, including recent climatic change and palaeoclimatic studies, were also discussed and will be highlighted here.

A new Southern Hemisphere data set, covering 1972 to 1979, has recently become available from the Australian Bureau of Meteorology and several aspects of the circulation down under were discussed, using these data. M.L. Blackmon and T.J. Phillips (National Center for Atmospheric Research (NCAR), Boulder, Colorado) examined the 1,000 mbar height data and showed that, between 1972 and 1979, there was a band of high variance between 40° and 60°S from south of Africa to New Zealand. Together with the results of K.E. Trenberth<sup>1</sup> which show a similar band of maximum variance in mid-tropospheric data at 500 mbar, these investigations suggest that the circulation of the Southern Hemisphere consists essentially of one major storm track. Corroborative evidence also came from J.C. Rogers (Ohio State University) who used Eigenvector analyses to show that negative correlations exist between high and middle latitudes in the same longitudinal band during the Southern Hemisphere summer. These results suggest that, when the westerlies in one sector of the hemisphere are strong, they are also strong in other sectors.

Chiu and Newell (MIT) described similar analyses for the Northern Hemisphere using upper air data for the period 1959–79 and concluded that the non-seasonal low-level changes during this period were dominated by the 'Greenland see-saw'

which has been described previously, notably by H. van Loon and J. Williams<sup>2</sup>. The see-saw operates when surface variables, pressure and temperature, in the region of southern Greenland and northern Scandinavia are negatively correlated. At higher levels (200 mbar) Chiu and Newell found that a persistent pattern, centred over Greenland, dominated. They suggested that this could be related to the Southern Oscillation which is a quasi-periodic four to six-year climate oscillation centred over the Pacific and linked to such important phenomena as El Niño.

First described as early as 1936 by Walker and Bliss, the Southern Oscillation is receiving a lot of attention because it is one of the best defined quasi-regular climatic events. Although primarily a tropical Pacific phenomenon, it has often been proposed as an important controlling mechanism for variations over much of the globe. Indeed, the studies described in San Diego confirmed old and elucidated new 'teleconnections' between the tropical Pacific and higher latitudes in both hemispheres. K.E. Trenberth and D.A. Paolino (University of Illinois) reported the variability of the Northern Hemisphere surface circulation showing connections with the Southern Oscillation. They used data from the period 1925–77 and Empirical Orthogonal Function (EOF) analyses to show that the dominant pattern in all seasons was one associated with anomalies centred on the Aleutian and Icelandic low-pressure areas. The time series of this pattern have significant periodicities at about six years which, they speculate, are associated with the Southern Oscillation.

Of the researchers who discussed synoptic analyses, only D.J. Shea (NCAR)

M.J. Salinger and C.B. Sear are at the Climatic Research Unit, University of East Anglia.

\*The first AMS conference on 'Climate Variations'. See footnote on previous page.

presented results from global data using EOF analyses. His sensitivity studies suggested that the first three EOFs of temperature were stable but that none of the EOFs of precipitation is so, implying a very high degree of randomness in spatial patterns of precipitation. This is an important consideration, particularly for those who are searching for a CO<sub>2</sub>-induced climatic change and are finding analogues from within the recent period.

Seemingly far removed from such studies were those using palaeoclimatic data to investigate long-term climatic change. Studies of deep-sea cores have provided considerable insight into climatic variations on such time scales. W.C. Prell (Brown University) concluded from examination of foraminifera assemblages from the western Arabian Sea that a distinctive assemblage occurs when the summer monsoon causes intense upwelling. The sediments provide a record of monsoonal upwelling variations over the past 150 thousand years which suggests that the monsoon was more intense from about 8,000 to 10,000 years ago — a time of increased precipitation in Africa and Arabia<sup>3</sup>. The more intense monsoon coincides with the dates of maximum Holocene warming reported in Australia and New Zealand. There is now mounting evidence that the most pluvial and warmest periods of the past 15,000 years did not occur simultaneously throughout the globe and this should be borne in mind by those attempting to produce analogues for a warmer, high-CO<sub>2</sub> world using data from the Holocene.

A model of oceanic moisture feedback to ice sheets in the Northern Hemisphere was advanced by W.L. Ruddiman and A. McIntyre (Lamont-Doherty Geological Observatory, Columbia University; see ref.4). They have found evidence from a core in the central North Atlantic that, for the past 250 thousand years, North Atlantic moisture feedback has amplified Milankovitch forcing, defined in summer insolation, from the 23,000-year precessional cycle, and have postulated a feedback mechanism in which orbitally controlled changes in summer insolation over the North American ice regulate meltwater/iceberg outflow to the North Atlantic. The feedback is such that when summer insolation is high, ice rafting into the ocean causes the formation of a low-salinity layer at the surface. This in turn restricts moisture transport to the ice sheet. The opposite happens when insolation is low. Accelerated shrinkage or growth

occur accordingly and positive feedback results.

Among the many modelling studies discussed were some concerned with palaeoclimatic circulation. A low-order general circulation model (GCM) of the Earth's atmosphere, developed by B.L. Otto-Blieser (University of Wisconsin-Madison), was used by J.E. Kutzbach (at the same institution) and T. Webb III (Brown University) to simulate the climate 18,000 years ago. The rationale behind the use of a low-resolution model was that palaeoclimatic data have similar resolution. If the model is successful, then both modern data and data from 18,000 BP can be used to simulate climate in the intervening epochs. Their simulation of the climate of 18,000 BP had an enhanced westerly gradient across the US with the strongest flow just south of the ice sheet, in agreement with earlier GCM results<sup>5</sup>.

S.L. Thompson, E.J. Barron and S.H. Schneider (NCAR) reported energy balance modelling studies, based on the work of Barron *et al.*<sup>6</sup>, which investigated the causes of the ice-free mid-Cretaceous climate. Their results suggest that changes in zonal palaeogeography (continental drift) are an important forcing mechanism for climate change but that meridional temperature distribution cannot be simulated successfully without additional feedback assumptions. Their model included albedo and cloud feedbacks and meridional heat transport changes but could still not model the Cretaceous temperature distribution. Thompson and his colleagues speculated that to attain Cretaceous levels of warming, the Earth could require additional external forcing from the Sun. Although indirect, this is one of the first quantitative investigations implicating solar constant changes as causes of climatic change on the time scale of hundreds of millions of years.

Few new results were presented on the recurrent theme of carbon dioxide, but modellers of a CO<sub>2</sub>-warmer Earth did not get away from San Diego without warnings, notably from S.H. Schneider and S.L. Thompson that a CO<sub>2</sub>-warmer Earth has so far been modelled using an instantaneous increase in CO<sub>2</sub> concentration, whereas, of course, in the real world the concentration increase occurs very slowly<sup>7</sup>. Thompson and Schneider suggest that the thermal inertia of the upper ocean, together with vertical mixing of the deep ocean waters, could delay the response of the globally averaged surface temperature to increasing CO<sub>2</sub> by a decade or so relative to equilibrium calculations. They suggest further that the evolution of surface climate anomalies, as CO<sub>2</sub> increases, may be significantly different from that suggested by equilibrium modelling results. They stress that the nature of this transient response is a major uncertainty in present CO<sub>2</sub>-climate change studies and that further research into this topic is urgently required. □

## The return of Halley

from David W. Hughes

HALLEY'S comet was last seen at the end of May 1911 when it was at about Jupiter's distance from the Sun and heading out into the Solar System beyond the orbit of Neptune. Aphelion was reached in 1948. The comet has been on its way back ever since.

The search for Comet Halley began, rather optimistically, in November 1977 when observers at two of the United States' largest telescopes unsuccessfully attempted to recover it more than eight years before the predicted perihelion passage date of 9 February 1986. The comet's estimated magnitude was fainter than 26 at that time.

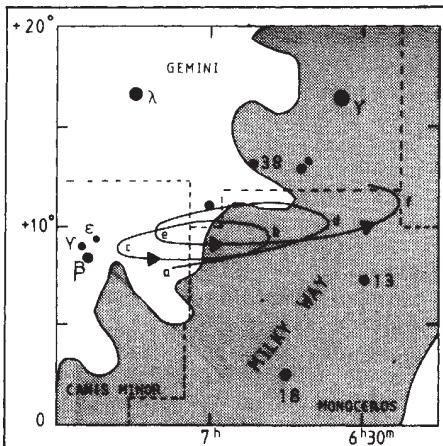
The early recovery of the comet is of considerable importance as this would enable a valuable check to be made on the predicted orbital parameters. Comets are much affected by the gravitational perturbations induced by the major planets. In the case of Halley this causes a slow variation in the period, between extreme values of about 76.1 yr and 79.4 yr, and a continuous increase in the longitude of the nodes. Spacecraft missions, like the planned Giotto Comet Halley flyby, require an accurate knowledge of the comet orbit to enable the spacecraft to fly as close to the comet as possible. At the last apparition the comet was recovered by Max Wolf using a 1 hour photographic exposure with the 72 cm Heidelberg reflecting telescope. The comet appeared as a small fuzzy image of magnitude 16. It was 3.4 AU from the Sun, in the outer regions of the Asteroid belt and 7.5 months from perihelion passage. At the present time the comet is 14 AU from the Sun and has an estimated magnitude of 24.7. (A change of 1 in magnitude is equivalent to a brightness ratio of about 2.5, so a comet of magnitude 16 is 3,000 times brighter than a comet of magnitude 24.7.)

The estimated path of Comet Halley is shown in the accompanying figure. Data for this figure have been extracted from *The Comet Halley Handbook: an Observer's Guide* which has been produced for the International Halley Watch by Donald K. Yeomans of the Jet Propulsion Laboratory, Pasadena, California (JPL publication 400-91 1/81). The comet is near the boundary of the constellations Canis Minor and Monoceros. The fact that the Milky Way acts as a backdrop obviously makes the search more difficult. The loopy path of the comet against the sky background is caused mainly by the movement of the observer. The comet is

David W. Hughes is in the Department of Physics, University of Sheffield.

1. Trenberth, K.E. *Mon. Weath. Rev.* (in the press).
2. van Loon, H. & Williams, J. *Mon. Weath. Rev.* **104**, 366 (1976).
3. Street & Grove *Nature* **261**, 385 (1976).
4. Ruddiman, W.L. & McIntyre, A. *Science* **204**, 173 (1979).
5. Williams, J. in *Climatic Change* (ed. Gribbin, J.) (Cambridge University Press, 1978).
6. Barron, E.J. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30**, 17 (1980).
7. Schneider, S.H. & Thompson, S.L. *J. geophys. Res.* (in the press).





The predicted path of Halley's comet for the next two years. The comet is mainly in the constellations Canis Minor and Monoceros. a, b, c, d, e and f represent its positions during January 1981, April 1981, October 1981, April 1982, October 1982 and April 1983 respectively. In January 1981 the comet had an estimated nuclear magnitude of +24.8 and this brightens steadily to about +22.5 in April 1983.

seen at maximum right ascension in October and at minimum right ascension in April. In this time interval the Earth has moved about 300,000,000 km, from one side of its orbit to the other. As the comet approaches the Sun the size of the loop increases. The most favourable time of the year for recovery will be when it has its maximum rate of angular movement against the sky background and can thus be more easily distinguished from a star. This occurs during the horizontal portions of the curve in the figure, at times roughly corresponding to winter and summer.

The fame of Halley's comet is not due to its brightness. Indeed, many of the accounts we hear from parents and grandparents of the appearance of the comet in 1910 do not refer to Halley at all but to the Daylight Comet, 1910 I, also known as the Great January Comet. Unfortunately, the forthcoming apparition will be unfavourable. Many of the waiting public will be disappointed as the comet will not be obvious to the naked eye. When Halley is at its most active it will be on the opposite side of the Sun to Earth and its maximum magnitude will be only about 2.9, about 2.5 times fainter than Polaris, the pole star.

Yeomans reviews the observing prospects and provides a series of charts showing the elevation, azimuth and magnitude of the comet in the morning and evening sky for latitudes 40°N, 30°N, 20°N, 20°S and 30°S. Two things are only too clear. Southern Hemisphere observers are better placed than Northern Hemisphere ones and people living at latitudes above 50°N are going to have considerable problems. It will be worth the effort though. One thing is certain, the similarity between our 'three score years and ten' and Halley's orbital period makes this a once in a lifetime experience for the large majority of us. □

## Hybrid myosins

from Dixie W. Frederiksen

A recent paper by Sellers, Chantler, and Szent-Gyorgyi<sup>1</sup> of Brandeis University has provided important new insight into the control of contraction in molluscan muscle and into the properties of myosins from a wide variety of species and cell types. It is the interaction of actin with myosin that brings about activation of the myosin Mg-ATPase and the production of mechanical work for contraction or cell motility<sup>2</sup>. The intracellular concentration of free calcium controls this interaction by mechanisms which differ in different systems.

In vertebrate striated muscle the primary effect of calcium is on troponin, a protein bound to actin and to tropomyosin on the thin filament of the sarcomere<sup>3</sup>. Conformational changes induced by calcium lead to a removal of inhibition of the actin-myosin interaction. In the striated muscles of molluscs and some other invertebrates, calcium acts directly on myosin in the thick filament to induce conformational changes that allow myosin to bind to actin<sup>4</sup>. In smooth muscle the situation is less well understood but the primary action of calcium appears to be an activation of a myosin light chain kinase. This produces a phosphomyosin which, unlike the unphosphorylated protein, can be activated by actin<sup>5</sup>. Additional effects of calcium have been reported for vas deferens and vascular smooth muscle<sup>6</sup> and it has also been shown that smooth muscle contraction may be calcium regulated through an actin-linked system<sup>7</sup>.

In all muscle and in most non-muscle cells, the myosin molecules<sup>8</sup> are hexameric species that consist of two 'heavy' (200,000 molecular weight) and four 'light' (14,000–31,000 molecular weight) polypeptide chains. There are two classes of light chains; the two 'non-essential', or 'regulatory', light chains are relatively easily removed from the major portion of myosin without loss of biological activity while the two 'essential' or 'alkali' light chains are removed from myosin only after rather harsh treatment with alkali or denaturant and their loss is accompanied by loss of ATPase activity. The intact myosin molecule is highly asymmetric with helical portions of the heavy chains contributing most to the length of the molecule. The light chains and globular portions of the heavy chains comprise a pair of structures at one end of the molecule that each contain an actin-binding and an ATPase site.

The detailed structures and functions of the myosin subunits vary from one cell type to another and, in their recent paper, Szent-Gyorgyi and his colleagues illustrate some important similarities and distinctions

among the various regulatory light chains from different myosins. Using some of the techniques they developed earlier<sup>9</sup>, they prepared approximately twenty different hybrid myosins containing the heavy and alkali light chains of scallop adductor myosin and two regulatory light chains from other myosins. They then tested regulatory light chains from vertebrate and invertebrate striated muscle, gizzard smooth muscle and platelets. Since all of these foreign subunits bind easily to light chain-depleted scallop myosin, the heavy chain-binding sites on all the regulatory light chains must be similar, as had been previously suggested by Weeds and his collaborators<sup>10</sup>.

The properties of the hybrid myosins are consistent with the idea that smooth muscle, as well as molluscan and some other invertebrate striated muscle myosins, are calcium-sensitive regulatory switches for the contractile process, but the vertebrate striated muscle has no myosin switch. Not surprisingly, the regulatory light chains from scallop and other molluscs restore full calcium ion sensitivity to the scallop myosin. In addition, light chains from crab, cricket, gizzard and platelets also form hybrids that bind calcium and require the divalent cation for actin activation of the ATPase. (The concentration of free calcium necessary for activation of these hybrids is higher than that required for the molluscan hybrids.) Vertebrate and lobster striated muscle myosin light chains produce calcium-insensitive hybrids without high-affinity calcium binding sites. Phosphorylation of the vertebrate striated and smooth muscle light chains (see *Myosin phosphorylation* *Nature* **287**, 191; 1980) did not affect the enzymatic properties of the hybrid myosins.

The Szent-Gyorgyi group has also resolved some of the puzzling questions about the mechanism of calcium control in molluscan muscle. Because the first regulatory light chain is much more easily removed from scallop myosin than the second, it seemed possible that the two light chain binding sites on scallop myosin are not identical. In an elegant series of experiments using labelled regulatory light chains, they showed that the light chains are removed in an order independent of the order of combination. The subunits bind to the main portion of myosin in a cooperative manner. The cooperativity is negative,

1. Sellers, Chantler & Szent-Gyorgyi *J. molec. Biol.* **143**, (1980).
2. Adelstein & Eisenberg *A. Rev. Biochem.* **49**, 921 (1980).
3. Weber & Murray *Physiol. Rev.* **63**, 612 (1973).
4. Szent-Gyorgyi *Biophys. J.* **15**, 707 (1975).
5. Dabrowska *et al.* *Biochem.* **17**, 253 (1978); Hathaway and Adelstein *Proc. natn. Acad. Sci. U.S.A.* **76**, 1653 (1979).
6. Chacko *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 129 (1977); Rees & Frederiksen *J. biol. Chem.* **256** (1981).
7. Mikawa *et al.* *J. Biochem. (Tokyo)* **82**, 1789 (1977).
8. Mannherz & Goody *A. Rev. Biochem.* **45**, 427 (1976).
9. Kendrick-Jones *et al.* *J. molec. Biol.* **104**, 747 (1976); Chantler & Szent-Gyorgyi *J. molec. Biol.* **138**, 473 (1980).
10. Weeds *et al.* in *Calcium Binding Proteins and Calcium Function* (eds Wasserman *et al.*) **222** (Elsevier North-Holland, Amsterdam, 1977).

Dixie W. Frederiksen is in the Department of Biochemistry, Vanderbilt University, Nashville, Tennessee.

however, since the affinity of myosin for light chain drops dramatically after the first one is bound. The affinity of light chain-depleted scallop myosin for foreign light chains is as great or greater than the

affinity for scallop light chains. Although this may be a fortuitous property of scallop myosin, it is intriguing to ask what might be learned using other light chain-depleted myosins in similar experiments. □

## Two distinct types of oceanic plateau

from Peter J. Smith

THE many refraction studies carried out over the oceans during the past twenty years have shown the basic structure of the oceanic crust to be remarkably uniform. The three-layer model proposed by Raitt (in *The Sea*, Wiley, 1963) in pre-plate tectonic, and almost pre-ocean floor spreading, days still generally holds good, notwithstanding the discovery of systematic variations in some lithospheric properties with age. At the same time, however, the ocean basins contain some obvious anomalous features, of which active ridges are the most conspicuous and hot spots the most controversial. And then there are the 'ocean plateaus' — broad aseismic platforms standing high above the ocean floor and usually topped with thick deposits of calcareous sediments.

Some of these plateaus (Rockall Bank and Seychelles Plateau, for example) have long been regarded as continental fragments, presumably broken off from larger landmasses during continental redistri-

bution. But others (for example, Broken Ridge and Shatsky Rise) are thought to be truly oceanic, although the processes by which they are formed are unknown. To test this two-fold classification and put it onto a more systematic basis, Carlson *et al.* (*Earth planet. Sci. Letts* **51**, 171; 1980) have now collated the seismic refraction data from 11 of the many oceanic plateaus known, showing that the two types are structurally distinct as might well be inferred from their supposedly different origins.

Of the 11 features studied, three were already known to be underlain by rock of continental type. Thus Rockall contains Lower Eocene granites; granitic rocks are exposed on the Seychelles Plateau; and gneissic rocks have been recovered from the basement of the Falkland Plateau by the Deep Sea Drilling Project. In addition,

Saya de Malha Bank is apparently part of the same physiographical feature as the Seychelles, the Mozambique Plateau is close to Africa and thought to be a piece broken from it, and Lord Howe Rise is likewise thought to be a fragment from nearby Australia, although in these three cases there is apparently no direct evidence of the presence of continental material.

But direct evidence or not, the seismic velocity-depth profiles suggest that these six platforms are of similar construction and thus members of a class. In three cases (Seychelles, Rockall, Lord Howe) upper mantle velocities ( $>8.0 \text{ km s}^{-1}$ ) are observed at about 30 km below sea level, and immediately above the Moho there is a 14–18 km thick lower crustal layer with seismic P wave velocities in the range  $6.0\text{--}7.0 \text{ km s}^{-1}$ . Unfortunately, the profiles from the three other plateaus (Mozambique, Saya de Malha, Falkland) do not extend down to the Moho, although the same basal crustal layer is present in Mozambique and Saya de Malha. Still more unfortunately, the Falkland profile does not even reach that far, but is consistent with the others in having a wide range of velocities of up to  $6.4 \text{ km s}^{-1}$  in its upper parts.

This upper-level velocity variability is also evident in eight seismic profiles from the five platforms suspected not to be of continental origin. But there the resemblance ends. Generally (five profiles — Shatsky, Broken Ridge, Fiji) the mantle is reached at depths of only 18–25 km, and immediately above the Moho there is a lower crustal layer 6–15 km thick having seismic velocities in the range

Peter J. Smith is the Editor of 'Open Earth'.



### THE FALLS OF NIAGARA IN WINTER

In the first week of last February it fell to my lot to make very hurriedly the transcontinental journey of 3500 miles from San Francisco to New York. Before starting I resolved that the one stoppage which I could allow myself *en route* should be made at Niagara. I was confirmed in my resolve by seeing statements in various American papers to the effect that, owing to the long-continued and exceptionally severe cold of the present winter, the Ice-mountains at the Falls were higher than had ever been previously known.

In the Niagara gorge numerous springs discharge themselves into the chasm at various points in the precipitous rocky sides, and at these points numerous collections of huge and massive icicles appeared as though adherent to the rock, measuring perhaps seventy or eighty feet in length, and eight to ten feet in irregular diameter.

The boulders in front of the American Fall are, as it were, gigantic nuclei, round which the frozen spray accumulates, and produces the

Ice-mountain of which we hear so much, and the remains of which are not unfrequently to be seen even by summer visitors. The average height of this is about half the total height of the Fall, but this winter it has attained to the unprecedented height of within twenty feet of the top of the Fall!

Within the last few years a considerable portion of "Table Rock" has fallen away. In its present condition a stream of water about one foot in thickness falls over it in summer, and, owing to the amount of its overhanging, it is easy to get between this Fall and the rock, and thus to be "behind Niagara." At the time of my visit (February 8th), however, the whole of this portion of the Fall was *completely frost-bound*. Enormous icicles, of the most surpassing beauty, depended from the rock above, while at my feet were masses of the frozen spray from the Horse-Shoe Fall. The intense emerald green of the water of that Fall, seen through and between these magnificent ice-pendants, could be reproduced by no artist, but will never be effaced from my memory. The accompanying woodcut, photographed on to the wood block from a photographic picture taken a few days prior to my visit, will, to those who know the place, give some faint idea of the beauty of the scene, and of the gigantic scale of the icicles.

From *Nature* **23**, 31 March, 511, 1881.



Gigantic icicle under Table Rock. The upper right-hand corner is rock; a portion of the Canadian (or Horse-Shoe Fall) is on the left. The whole of the apparent ground is a mass of frozen spray which has accumulated many feet in thickness on the shingle at the foot of the rock.



7.3–7.6 km s<sup>-1</sup>. Then above that there is a unit with typical oceanic layer 3 velocities. The only exception to all this (apart from the Manihiki Plateau whose profile is too short to indicate much) is the Ontong–Java Plateau, the two profiles of which demonstrate a curious mixture of both ‘continental’ and oceanic characteristics. On the face of it, this spoils the story rather, although it could quite genuinely represent a composite structure with both continental and oceanic elements.

Leaving that anomaly aside, however, there are clearly just two basic and distinct structural patterns, suggesting two separate modes of origin. The two types of structure also correlate well with their respective counterparts elsewhere. Thus the first (continental) pattern is very similar to the seismic velocity–depth structure of continental shields, whereas the second (oceanic) is not. Then, the high basal

velocities evident in the second pattern have also been observed at the base of normal crusts in parts of the Pacific and the Atlantic, as well as in some ophiolites, which in fact strengthens the view that oceanic plateaus of the second kind are genuinely of marine origin.

Carlson and his colleagues admit that the number of plateaus for which velocity profiles are available is small and therefore regard their analysis as tentative, although the agreement between the seismic data and prior suppositions about two different origins inspires confidence. The mode of formation of the non-continental plateaus remains open, however. The new analysis seems to confirm the conclusion by Hussong and his colleagues (*J. geophys. Res.* **84**, 6003; 1979) that the truly marine plateaus are expanded sections of normal oceanic crust, but the cause of the expansion is a mystery. □

lymphocytes from many long-standing MG patients make anti-AChR in culture, it seems that the thymus is not the primary site of anti-AChR synthesis (J. Newsom-Davis, Royal Free Hospital, London). Thymic lymphocyte preparations (irradiated to destroy B cells) can also enhance the *in vitro* anti-AChR synthesis by MG peripheral blood lymphocytes, possibly due either to antigenic stimulation by AChR on thymic cells, or to the presence of a thymic helper-T cell population specific for anti-AChR (or both) (Newsom-Davis). Other cell types may also be involved, however, as N. Shinomiya (Tokyo Medical and Dental School) showed that normal, but not MG, peripheral blood T cells inhibited anti-AChR synthesis by MG lymphocytes. This suggests the absence of a specific suppressor-T cell in MG subjects.

This latter finding indicates that even normal subjects may have AChR-sensitive lymphocytes and display low levels of anti-AChR. Anti-AChR has been detected in the sera of 18 per cent of MG relatives (R. Pirskanen, Karolinska Institute) although it did not correlate with HLA-B8, the HLA antigen most closely associated with MG, nor with the electromyographic abnormality also found in a proportion of MG relatives. Others have demonstrated very low levels of anti-AChR in normal subjects (T. Mittag, Mount Sinai School of Medicine), in cultures of normal lymphocytes with MG T cells (R. Kelley, UCLA School of Medicine) and in purified Torpedo AChR (Shinomiya, Tokyo Medical and Dental School). The existence of sensitized B lymphocytes would not be surprising if the normal immune system was exposed to AChR, and it may be relevant that AChR-like molecules have been detected on normal peripheral lymphocytes (G. Morrel, Wayne State University; D. Richman, University of Chicago) and mouse thymic lymphocytes (S. Fuchs, Weizmann Institute) although  $\alpha$ -bungarotoxin binding has not been demonstrated.

Work on EAMG may help to elucidate the involvement of genetic factors. When animals are immunized against AChR, a proportion of the anti-AChR antibodies cross-react with the animals’ endplate AChR, resulting in weakness and neuromuscular block. Not all species or strains respond equally and P. Berman (Salk Institute) showed that susceptibility to clinical EAMG in inbred strains of mice was not related to the anti-AChR titre, its ability to increase AChR degradation, or a difference in the safety factor for transmission at endplates from different genetic strains. The greatest susceptibility was found in those with a particular major histocompatibility complex (H-2<sup>b</sup>) and immunoglobulin (IgI<sup>b</sup>) genotype. The involvement of the IgC<sub>H</sub> region suggests that overt myasthenia in mice depends on particular polymorphic genes coding for the constant region of immunoglobulin

## Idiotypic restriction in myasthenia gravis antibodies

from Angela Vincent

It is now reasonably certain that anti-acetylcholine receptor antibodies (anti-AChR) play a pathogenic role in the reduction in endplate acetylcholine receptors (AChR) that underlies the defective neuromuscular transmission characteristic of myasthenia gravis (MG). Experimental autoimmune MG (EAMG) in animals immunized against purified AChR, human anti-AChR in over 80% of MG patients, and passive transfer of MG to mice using MG immunoglobulins had all been reported by 1975 and progress made since then was the subject of a recent conference\*.

It has been shown that, although in any individual the initial anti-AChR titre gives no indication of the severity of the disease, subsequent anti-AChR levels tend to correlate with changes in the clinical state. Various mechanisms thought to be involved in loss of AChR function have been demonstrated, including complement-dependent damage to the postsynaptic membrane (A. Engel, Mayo Clinic), increased AChR degradation caused by cross-linking of AChR by divalent antibody (D. Drachman, Johns Hopkins University), and direct ‘block’ of AChR function by binding of antibody to the ACh/ $\alpha$ -bungarotoxin binding site (D. Drachman; B. Fulpus, University of Geneva). Antibodies can also modify rather than block AChR function, as E. Albuquerque (University of Maryland)

showed that a factor in MG serum can slightly increase the open lifetime of the individual AChR-associated ion channels. This effect could only be seen, however, when the serum had been applied in the presence of histrionicotoxin, a toxin that reversibly blocks the channel in its open state.

Although the pathogenesis of the disease appears to be established, the fundamental aetiology and cellular immune mechanisms are not. Is there a precipitating event such as virus infection, and how are genetic factors involved? Is there a specific lymphocyte abnormality or a failure of normal control mechanisms? Any answer to these questions must include an explanation of the role of the thymus. Many patients have ‘thymitis’ with germinal centres, and some have a thymic tumour. MG symptoms, particularly in the former group, often respond well to thymectomy although improvement may take months or years. One possibility is that the thymus is the site of both the antigen and the initiating event. It has been recognized for several years that cultured thymic epithelial cells can display muscle-like antigens including AChRs (K. Wekerle, Max Planck Institute). Wekerle proposed that helper-T cells were sensitized to AChR in the thymus and then emigrated to the periphery where anti-AChR synthesis took place. This may well be the case because, although thymic

Angela Vincent is in the Department of Neurological Science, The Royal Free Hospital, London.

\*The sixth conference on myasthenia gravis was held in New York from December 2nd to 4th, 1980, organized by Dr D. Grob (Maimonides Medical Center) and sponsored by the New York Academy of Sciences and the Myasthenia Gravis Foundation.

heavy chains. A similar association (with the human Gm allotypes) has been reported in MG (Nakao *et al. Lancet*, 677; 1980). These findings might explain the generally poor correlation between the total anti-AChR titre and disease severity. Clearly, specific subpopulations of antibodies may also be important in MG and will be defined when human monoclonal anti-AChR is available.

In EAMG several workers have used monoclonal antibodies produced by hybridomas to analyse the anti-AChR. A single monoclonal antibody reactive with mammalian AChR can passively transfer EAMG to mice (J. Lindstrom, Salk Institute) or rats (V. Lennon, Mayo Clinic). Surprisingly, antibodies raised in inbred rats against each of three distinct monoclonal antibodies inhibited the activity of the others suggesting that they shared a common idio type (the antigen-specific, hypervariable region) (V. Lennon). However, each anti-idiotypic antibody inhibited only a proportion of the anti-AChR in polyclonal EAMG sera and

when anti-idiotypic-producing animals were immunized directly against AChR they still developed electrophysiological evidence of EAMG.

J. Lindstrom has used monoclonal antibodies raised against intact AChR and its polypeptide subunits to map determinants on receptor extracted from many different species. The main immunogenic region appears to be on the  $\alpha$ -subunit of Torpedo and eel AChR and similar determinants are found on AChR from mammals including humans. Three monoclonal antibodies reacting with  $\alpha$ ,  $\beta$  and  $\gamma$  subunits can bind to human AChR and inhibit the binding of a proportion of human anti-AChR (S. Tzartos, Salk Institute). These results are important because they suggest that the determinants to which human anti-AChR bind are more restricted than was previously inferred from the heterogeneous characteristics of MG anti-AChR (as defined by variability in  $\kappa$  and  $\lambda$  light chains, binding to the  $\alpha$ -bungarotoxin binding site, cross-reactivity with other AChR preparations).

Moreover, using a polyclonal anti-idiotypic serum raised in rabbits against human anti-AChR, A. K. Lefvert (Karolinska Institute) has found some sharing of anti-AChR idiotypes between different MG patients.

The treatment of MG has improved since the autoimmune basis for the disease has been widely recognized. Immunosuppressants are used much more freely with good results, and plasma exchange can produce marked temporary improvement. Although several interesting experimental approaches to treatment, including immunization against modified AChR and anti-idiotypic induction in rabbits with EAMG (S. Fuchs), have met with some success, it will be some time before any can be tried in the human disease. Nevertheless, if there is less heterogeneity and greater idio type restriction in MG anti-AChR than previously thought, and if certain idiotypes can be shown to play a crucial role in the disease process, specific immunotherapy may become a realistic goal. □

## Disposing of dioxins by oxidation

from Alastair Hay

DISPOSAL of waste contaminated by polychlorinated dibenzodioxins (PCDDs) has long been a problem for sections of the chemical industry. Toxic and often remarkably stable, the PCDDs are produced as unwanted by-products during the manufacture of chlorinated phenols.

The best known and most extensively studied PCDD is 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). It is produced during the manufacture of 2,4,5-trichlorophenol — the starting point for the widely used herbicide 2,4,5-trichlorophenoxyacetic acid (2,4,5-T).

Although pure TCDD is very stable (a temperature in excess of 800°C is required for its thermal degradation), if dissolved in organic solvents such as methanol or benzene it is exceptionally unstable in the presence of ultraviolet light. The solvents act as hydrogen donors, replacing aromatic ring chlorines with hydrogen as TCDD is photolytically degraded. The reaction sequence proceeds in stages with only one chlorine atom being replaced at any one time. Eventually TCDD is degraded to the unchlorinated parent dibenzodioxin (D.G. Crosby *et al. Science* 173, 748; 1971). Many orders of magnitude less toxic than TCDD, dibenzodioxin too is decomposed by light (J.R. Plimmer *et al. Adv. Chem.* 120, 44; 1973).

The higher chlorinated dibenzodioxins, hexa-, hepta- and octachlorodibenzo-*p*-dioxin, produced in the manufacture of the wood preservative pentachlorophenol, are also subject to photolytic degradation (C. Rappe *et al. Ann. N.Y. Acad. Sci.* 320, 1;

1979).

Under field conditions TCDD can be subjected to photolytic degradation by sunlight when a suitable solvent/hydrogen donor is present. When a preparation of the herbicide 'Agent Orange' — a formulation of 2,4,5-T containing 15 p.p.m. TCDD — was exposed to sunlight, the dioxin was rapidly degraded with less than half remaining after 6 hours (A.S. Wong and D.E. Crosby in *Dioxin: Toxicological and Chemical Aspects* (eds Cattabeni, F. *et al.*, Spectrum, New York; 1978). When TCDD was dissolved in an oil/emulsifier mixture and exposed to sunlight, complete degradation occurred within 2 days (A.S. Wong and C.G. Crosby *op. cit.*).

In industry, photodegradation of PCDDs is not very practical and most manufacturers dispose of the chemicals by high-temperature incineration. A method of degradation using chemical processes might provide a more attractive option and in this issue of *Nature* (p.323), D.C. Ayres reports that PCDDs can be degraded efficiently by an oxidative procedure involving ruthenium tetroxide.

Ruthenium tetroxide is a powerful oxidizing agent which reacts with a wide range of organic substances, and according to Ayres, it can be used in solution in water or in organic solvents with no nucleophilic character, such as chloroform, nitro-

methane or carbon tetrachloride. Ayres reports that 2,7-dichlorodibenzo-*p*-dioxin (DCDD) was destroyed by ruthenium tetroxide in carbon tetrachloride solution when the cocktail was allowed to stand overnight at ambient temperature. The rate of oxidation increases with increase in temperature: at 30°C DCDD has a half life of 215 min whilst at 50°C, the half life falls to 38 min. The oxidation of TCDD is a slower process: at 20°C this chemical has a half life of 560 min in the ruthenium tetroxide mixture. At 70°C, however, the half life plummets to a mere 15 min.

Ayres believes that his results demonstrate that ruthenium tetroxide can be safely used for the detoxification of glassware and the periodical purging of industrial reactors to counteract the accumulation of PCDD residues. The process need not be expensive, Ayres claims, for ruthenium tetroxide need only be used in catalytic amounts after which it can be regenerated using a consumable secondary oxidant, such as hypochlorite.

Although Ayres does not know the nature of the fragments formed during the oxidation of the PCDDs, he reports that chlorophenols subjected to ruthenium oxidation undergo extensive decomposition to yield chloride ions and no significant levels of organic products. It remains to be seen whether this is the fate of the PCDDs, but it is clear that oxidation must now challenge incineration as the method of choice for disposing of unwanted polychlorinated dibenzo-*p*-dioxins in industry.

Alastair Hay is in the Department of Chemical Pathology at Leeds University.



## ARTICLES

## Orbital signature of interglacials

G. Kukla\*, A. Berger†, R. Lotti\* &amp; J. Brown\*

\* Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

† Institute of Astronomy and Geophysics, Université Catholique de Louvain, Louvain-La-Neuve, Belgium

*A specific orbital configuration—high obliquity combined with the June perihelion—marked the beginning of the past three interglacials. This suggests that the primary cause of the glacial cycle may be astronomical. An astronomical climate index (ACLIN) is introduced which combines the three orbital variables in the time-lag bivariate model designed to predict the major climate changes in the late and middle Pleistocene, and in the near future. ACLIN closely correlates with the major climatic events revealed by independently dated proxy climate indicators of the past 130,000 yr. It successfully differentiates the interglacials, and displays a 100,000-yr periodicity. It predicts an early end of the present interglacial and the start of a new one in 114,000 yr.*

SEVERAL recent observations have shown that the principal periods in the radiometrically dated palaeoclimatic records match in length the periodicities of the Earth's orbit<sup>1-5</sup>. These two systems are probably linked by a cause and effect relationship. Although there have been several attempts to explain the palaeoclimatic record in terms of insolation input, none of the selected palaeoinsolation curves correlates with the different amplitude of interglacials and interstadials. The models simulate the exceptional character of the post-glacial warm interval only when the nonlinear terrestrial response to insolation forcing is included. Different rates of ice growth and decay, a nonlinear isostatic adjustment to the ice-load, thermal memory of the oceans, and so on, represent the terrestrial input<sup>6,7</sup> (for a review, see ref. 8).

It has been argued, however, that the internal dynamics of the ice sheets can explain the dominant 100,000-yr periodicity of palaeoclimatic records<sup>9,10</sup> and consequently, that the astronomical modulation of the Earth's climate has, at best, only a secondary role.

We have attempted to test whether the non-linear terrestrial response is necessary to explain the glacial-interglacial cycle. Is there an orbital configuration specific to an interglacial and can the exceptionally warm interglacial climate be explained by orbital configuration alone?

Interglacials are the key climatostratigraphic units recognized in Pleistocene strata worldwide—they reappear about every 10<sup>5</sup> yr. It is the interglacial environment in which mankind flourished. Any useful prediction of future natural climates has to be able to differentiate between the interglacials and the glacials. If a specific orbital signature of an interglacial was found it would provide: (1) a better understanding of climate; (2) a definition of predictors of future gross climate shifts; and (3) improvement in the dating of past climates.

In this article we first summarize radiometrically dated evidence of climatic changes during the past 350,000 yr; then

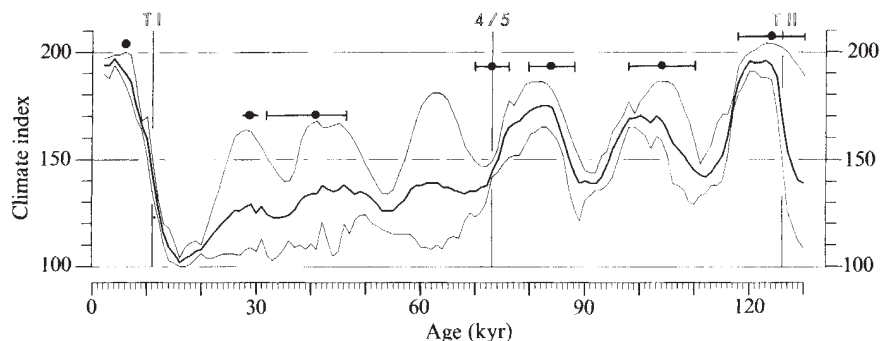
compare this record with the history of orbital perturbations and identify orbital configurations, if any, which correlate with interglacials, interstadials and stadials; and finally propose and test an index, based solely on orbital variables, with possible use as a predictor of past and future gross climate states.

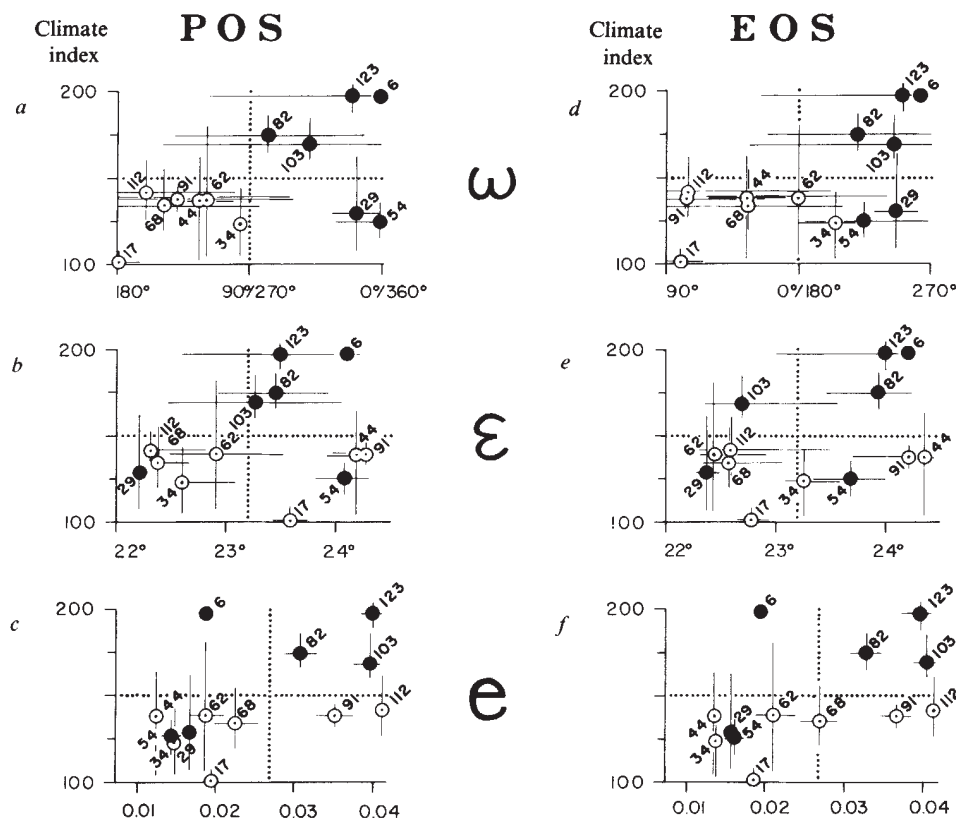
## Climatic variations of the past 350,000 yr

Past climates were reconstructed from indirect evidence on the size and volume of former ice sheets and on the change in habitats of flora and fauna<sup>11,12</sup>. Radiometric dates, counts of annual layers and interpolated sedimentation rates provided the time scale. Ocean records are the most complete<sup>13,14</sup>; land records are the most detailed and the best dated<sup>15</sup>. The accuracy of the dating decreases with time. Dating of the past 30,000 yr relies on radiocarbon; it is accurate to within  $\sim \pm 1,000$  yr; in the preceding 10,000 yr to within about  $\pm 5,000$  yr. Previous to that, because of contamination problems<sup>16</sup>, <sup>14</sup>C ages have to be considered minimum values. U/Th isotopic ages give the highest accuracy ( $\pm 5,000$ – $6,000$  yr) in the interval between 60,000 and 130,000 yr BP.

Figure 1 summarizes the climate records of the past 130,000 yr. It is based on three sets of proxy climate indicators dated by radiocarbon up to 50,000 yr BP and by U/Th decay series thereafter. The three groups are: (1) Pollen data from Grande Pile, France<sup>17-20</sup> and Tenaghi Phillipon, Greece<sup>21,22</sup>, transferred onto a scale where 100 units correspond to the peak glacial (between 15,000 and 20,000 yr BP) and 200 units to the peak present interglacial (6,000 yr BP). (2) Sea level estimates from dated coral reefs in Barbados<sup>23-30</sup>, Curacao<sup>24</sup>, Bermuda<sup>31,32</sup>, New Guinea<sup>33-36</sup>, Japan<sup>37</sup>, Loyalty Islands<sup>38</sup>, Australia<sup>39</sup>, Eniwetok Islands<sup>40</sup>, Florida<sup>24,41</sup> and the Bahamas<sup>41,42</sup>, and <sup>14</sup>C dated palaeo-sea levels in North Carolina<sup>43,44</sup>, Scotland<sup>45</sup>, and Australia<sup>46</sup>, transferred on the same relative scale. (3) Delta <sup>18</sup>O records<sup>3,47-49</sup> in cores RC11-120,

**Fig. 1** Climate severity index for the past 130,000 yr based on the three sets of proxy climate indicators (pollen, sea level and  $\delta^{18}\text{O}$ ), expressed on a relative scale where 100 units correspond to the peak glacial and 200 to the peak of the present interglacial. Input data and procedure are explained in text. Horizontal bars show the dating accuracy whereas the envelope gives the range of the climate index values. Note large uncertainty between 20,000 and 70,000 yr BP.





**Fig. 2** Orbital signature at the peak (POS) and at the onset (EOS) of major warm and cold episodes of the past 130,000 yr, dated by radioisotopes. The episodes are identified by the age of their peak, given in thousands of years. Vertical scale: climate severity index from Fig. 1. Horizontal scales: longitude of the perihelion ( $\omega$ ) in degrees, obliquity ( $\epsilon$ ) in degrees, and the eccentricity ( $e$ ). ● mark the episodes with the POS  $\omega$  from 270° to 360° to 90°. Discussion in text.

METEOR 12392-1, V19-29, and V29-179 dated by correlation with sets 1 and 2 and by Toba Ash<sup>50</sup>, and plotted along the same scale. The index was averaged arithmetically, first within each group and second among groups 1, 2 and 3. In the pollen record from Grande Pile<sup>17-20</sup> in France, and Tenaghi Philippon<sup>21,22</sup> in Greece, the climate index  $\gamma$  was computed after the empirical formula:

$$\gamma = 4x + 2y + z$$

where  $x$ ,  $y$  and  $z$  are the warm, cool, and cold principal pollen producing species, respectively, taken as a per cent of the total pollen count. Here  $x$  is the sum of oak, fir, hornbeam, alder and hazel; for Grande Pile yew and beach were added; for Tenaghi Philippon Mediterranean hornbeam and pistachio were added.  $y$  is the sum of pine, birch and juniper; with spruce added for France, and willow added for Greece.  $z$  is grass in France, and the sum of sagebrush and herbs in Greece.

The pollen formula was selected after multiple comparisons of the results with published palaeotemperature reconstructions of the past 130,000 yr (ref. 51). The above formula gave the closest agreement. Details of the procedures will be described elsewhere.

The smoothing process removed the relatively sharp boundaries observed in the land based records<sup>52,53</sup> separating the cool and the temperate stages. The time series of most proxy climate indicators when inspected on a sufficiently detailed time scale, have a step function form<sup>53</sup>. They show relatively steady intervals several millenia long, interspersed with short episodes of relatively rapid change. The more recent of these shifts were dated close to 10,000 yr BP (ref. 54), 23,000 yr BP (ref. 55), and 32,000 yr BP (ref. 55). Some such records may result from a step response of a proxy climate indicator with a broad climatic tolerance to a smooth continuous change of climate. However, recent data indicate that the climate over large regions changed in steps<sup>52,56</sup> along a generally sinusoidal path.

The principal features of the palaeoclimatic history of the past 130,000 yr (shown in Fig. 1) can be summarized as follows. The past 10,000 yr and the interval between ~115,000 and 127,000 yr BP had an interglacial climate, considerably warmer than the rest of the time. There were only two interglacials in the

past 130,000 yr, both began by a pronounced stepwise environmental change accompanying a full scale deglaciation (termination). The most severe part of the last glacial lasted from ~22,000 to ~13,500 yr BP (ref. 57) and continental ice sheets were most extensive between 19,000 and 15,000 yr BP (ref. 58). At that time the sea level was 80 m (ref. 44), and possibly 130 m (ref. 43), lower than today. Similar environmental conditions and, by inference, climate, marked the next-to-last full glacial between 130,000 and 140,000 yr BP (ref. 15). Milder inter-

**Table 1** Obliquity and June insolation at the start of the mild and cool intervals

| Mild intervals |            |                                                        | Cool intervals |            |                                                        |
|----------------|------------|--------------------------------------------------------|----------------|------------|--------------------------------------------------------|
| Time (kyr)     | $\epsilon$ | June solstice insolation 90° N (Ly day <sup>-1</sup> ) | Time (kyr)     | $\epsilon$ | June solstice insolation 90° N (Ly day <sup>-1</sup> ) |
| +10            | 22.61°     | 1,105                                                  | +19            | 23.16°     | 1,091                                                  |
| -11*           | 24.20°†    | 1,197                                                  | -1             | 23.58°     | 1,086                                                  |
| -34            | 22.60°     | 1,112                                                  | -22*           | 22.79°‡    | 1,048‡                                                 |
| -60            | 23.22°     | 1,146                                                  | -47            | 24.42°     | 1,132                                                  |
| -83            | 23.58°     | 1,197                                                  | -72            | 22.38°‡    | 1,018‡                                                 |
| -105           | 22.96°     | 1,185                                                  | -94            | 24.26°     | 1,075                                                  |
| -127*          | 24.04°†    | 1,239                                                  | -116           | 22.49°‡    | 991‡                                                   |
| -151           | 22.48°     | 1,132                                                  | -139*          | 23.57°     | 1,050                                                  |
| -176           | 23.75°     | 1,213                                                  | -164           | 23.87°     | 1,079                                                  |
| -198           | 22.85°     | 1,205                                                  | -187           | 22.62°‡    | 991‡                                                   |
| -220*          | 23.78°†    | 1,247                                                  | -209           | 24.29°     | 1,047‡                                                 |
| -242           | 23.25°     | 1,186                                                  | -231*          | 22.08°‡    | 972‡                                                   |
| -267           | 22.80°     | 1,129                                                  | -254           | 24.40°     | 1,104                                                  |
| -292           | 23.98°†    | 1,217                                                  | -280           | 22.90°     | 1,044‡                                                 |
| -314           | 22.54°     | 1,167                                                  | -303           | 23.52°     | 1,040‡                                                 |
| -335*          | 24.21°†    | 1,225                                                  | -324           | 23.36°     | 1,035‡                                                 |
| -355           | 22.36°     | 1,094                                                  | -345*          | 22.97°     | 1,049                                                  |

\* Warm interval following or cool interval preceding major deglaciation.

† Values equal or higher than that corresponding to a dated interglacial.

‡ Values equal or lower than that corresponding to the last peak glacial.



**Table 2** ACLIN 1 cutoff values

| Climate state            | ACLIN 1 peak     | Duration (kyr) |
|--------------------------|------------------|----------------|
| Interglacial             | $\geq 4.3$       | 10             |
| Temperate interstadial   | $\geq 3.5 < 4.3$ | 8              |
| Interstadial             | $\geq 2.5 < 3.5$ | 4              |
| Stadial and protostadial | $< 2.5$          |                |

stadials punctuated the glacial between 115,000 and 22,000 yr BP: the first two episodes are sometimes referred to as temperate interstadials. The environment then locally approached, but globally did not reach, interglacial conditions. Remaining interstadials were cold with arctic climate in the northern middle latitudes. At  $\sim 73,000$  yr BP (ref. 14) the climate deteriorated sharply.

As seen in Fig. 1, the climate severity between  $\sim 70,000$  and 20,000 yr ago is relatively uncertain. While the deep-sea record and the pollen assemblages in Europe testify to harsh conditions and suggest a large volume of land based ice, the sea level estimates based on raised coral reefs imply a much smaller ice volume.

Palaeoclimates of the interval 130,000–350,000 yr BP are known in much lesser detail and are less accurately dated (Fig. 3). Four interglacials occurred in the past 350,000 yr; each was preceded by one of the major deglaciation dated to  $\sim 11,000$ , 127,000, 220,000 and 330,000 yr BP. The oldest date is based entirely on interpolated thicknesses of the Brunhes sediments in deep sea cores (ref. 15). Published estimates differ by about  $\pm 10\%$ . Continuous pollen bearing deposits<sup>51</sup>, soil-loess sequences<sup>15</sup>, and  $^{18}\text{O}$  series<sup>59</sup> show that the frequency and amplitude of climate oscillations before and after 130,000 yr BP was approximately the same.

## Climate index compared with orbital variables

The Earth circumsolar orbit is defined by three parameters:

- (1) The tilt  $\epsilon$  of the Earth axis to the plane of the ecliptic (obliquity). It varies from  $\sim 22^\circ$  to  $\sim 24.5^\circ$  within a basic period of  $\sim 41,000$  yr. With higher tilt, more insolation is directed to the high latitudes of both hemispheres in summer at the expense of low latitudes.
- (2) The longitude of the perihelion  $\omega$ . This is measured as the angular distance of perihelion from the autumnal equinox, and defines the season in which the Earth makes its closest approach to the Sun. At perihelion, the intensity of solar radiation reaching the Earth is at a maximum. The precession of the equinoxes, as the process is called, has an irregular periodicity of 13,000–25,000 yr with periods of  $\sim 23,000$  yr and 19,000 yr being the most common<sup>4</sup>.
- (3) The eccentricity  $e$  is the measure of the Sun's departure from the geometric centre of the Earth's orbit. High eccentricity increases the effect of  $\omega$  which varies from close to 0 to  $\sim 0.06$ , with periods of  $\sim 100,000$  yr and 410,000 yr.

The three orbital parameters at any past or future time can be computed from the equations of planetary mechanics<sup>68</sup>. Although we used recently revised calculations of Berger<sup>60,61</sup>, the calculations of others agree to about 350,000 yr BP but they diverge further back in time<sup>60</sup>. The precision of the calculations decreases with time due to the limited number of terms used in the expansions and to imprecise knowledge of the input parameters.

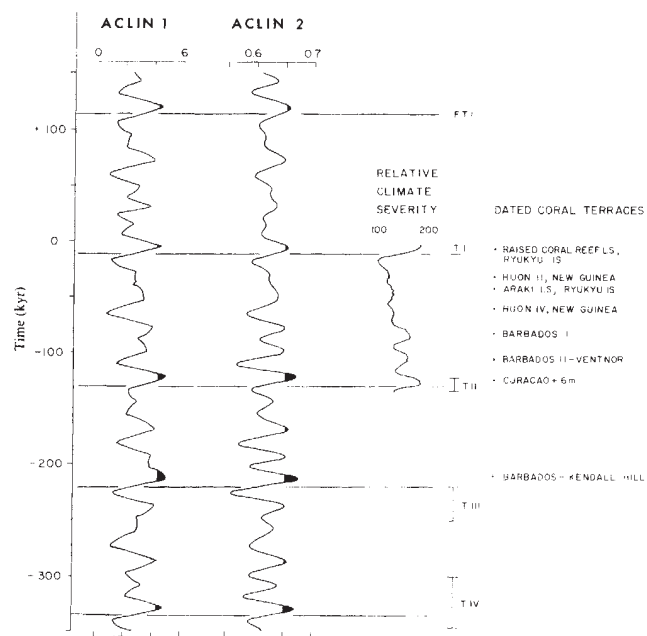
Figure 1 gives the timing of the major warm and cold climatic episodes with a wavelength of at least 10,000 yr. It also defines a starting (entry) date of each episode, half way between the neighbouring highs and lows. The orbital signature corresponding to the time of climate peaks, the peak orbital signature (POS), is plotted in Fig. 2a–c. Characteristics of the start of each episode, the entry orbital signature (EOS), is plotted in Fig. 2d–f. Each part of the figure is divided into quadrants defined by the mean values of the climate index and of the corresponding orbital measure.

Figure 2a shows, that mild or warm climatic episodes tend to peak when the longitude of perihelion  $\omega$  is close to  $0^\circ/360^\circ$  (September perihelion), whereas cool episodes culminate near  $180^\circ$  (March perihelion). We therefore label episodes with peak  $\omega$  between  $90^\circ$  and  $270^\circ$  by open circles and the remainder by full circles. This convention applies throughout Fig. 2. Figure 2b shows that warm climates tend to be associated with high obliquity, and cool climates with low obliquity. Finally, (Fig. 2c) cold climates seem to be associated with low eccentricity.

None of the warm episodes display simultaneously high values of eccentricity, obliquity and autumn perihelion. For instance, the present interglacial peak at 6,000 yr BP has low eccentricity, the 29,000 yr BP interstadial peak has a low obliquity and eccentricity, and the relatively cold episode at 91,000 yr BP has a high obliquity and eccentricity, but a longitude of the perihelion close to  $180^\circ$ .

The orbital configuration disagrees only once with the corresponding climate, at 54,000 yr BP, a time of high obliquity and of autumn perihelion. Although the eccentricity is low, it is not much lower than at the peak of the present interglacial. Why, then, should 54,000 yr BP mark a cold climate? One possibility is that the disagreement stems from poor dating and possible incompleteness of the sea level curve between 40,000 and 65,000 yr BP, which strongly influenced the average index (Fig. 1). Contrary to the master curve, the 54,000-yr BP interval shows a relative amelioration in both the pollen data and the  $\delta^{18}\text{O}$  record. However, as no marine terrace of such age is reported, we assume a low sea level between 45,000 and 60,000 yr BP.

Figure 2d–f shows the entry orbital signature as a function of the peak climate index. A phase shift of  $\sim 5,000$ – $6,000$  yr exists between the plotted orbital variables and the climate index. The mild and warm episodes show  $\omega$  close to  $270^\circ$  (June perihelion) whereas cool episodes close to  $90^\circ$  (December perihelion, Fig. 2d). Obliquity values are high during warm episodes and low during cold ones, with the same exceptions as in Fig. 2b. The eccentricity association is similar to Fig. 2c.



**Fig. 3** Plots of ACLIN 1 and 2 for the past 350,000 yr and future 150,000 yr. Higher index value correlates with warmer climate. ACLIN 1 interglacial peaks  $\geq 4.3$  and ACLIN 2 peaks  $\geq 0.654$  filled in black. Climate severity index for the past 150,000 yr same as in Fig. 1. Mean ages of the U/Th dated coral reefs<sup>23,24,36,37</sup> shown by dots. Vertical bars show the range of age estimates of terminations (TI–TIV) as obtained from deep sea cores<sup>15</sup>. FT I is the first expected future termination. Peaks as high or higher than that of 6,000 yr BP shown by arrow.

We conclude from Fig. 2 that the two interglacials, peaking at 6,000 yr BP and 123,000 yr BP, possess specific orbital characteristics which are absent during the remaining periods. They start at a time of June perihelion and reach a climax at a time of September perihelion, with high obliquity throughout the first half of the interval. Only the interstadial at 82,000 yr BP shows a similar, though markedly weaker character.

To test the correlation of high EOS obliquity with interglacials, all June perihelions ( $\omega = 270^\circ$ ) during the past 360,000 yr BP were listed in Table 1, together with the corresponding obliquity. The known four interglacials started at 335,000, 220,000, 127,000 and 11,000 yr BP. The interglacial starting at 220,000 yr BP shows the lowest EOS obliquity of  $23.8^\circ$ , all three remaining interglacials start with higher values. All interstadials except the one at 292,000 yr BP have an obliquity lower than  $23.8^\circ$ . Thus, at least within the past 250,000 yr, an EOS obliquity of  $\geq 23.8^\circ$  together with a June perihelion, is an exclusive characteristic of interglacials. However, this observation does not necessarily imply the existence of a physical mechanism directly linking the two observed properties with terrestrial climates. Some other variable indirectly related in space and time may provide the critical solar-terrestrial link.

It might be expected that the combined impact of high obliquity and of a June perihelion be fully reflected in the insolation values at high latitudes and that a single insolation curve for a selected latitude could serve as a predictor of interglacials. Table 1 shows that this is not the case. Even at the North Pole, where the departures are largest, the insolation at the onset of the postglacial was not any higher than during the four earlier interstadials. The physical explanation of the climatic importance of the high tilt probably lies in the dynamics of the seasonal sea-ice covers. These disintegrate in the late spring and summer as a result of high solar radiation<sup>62,63</sup>. Only the obliquity cycle, but not the precessional cycle, increases the insolation of the pack ice fields of both hemispheres in phase.

The special importance of the autumn perihelion is linked to the dynamics of snow fields. The seasonal advance of snow fields in autumn has been closely correlated with the seasonal change of insolation<sup>64</sup>, thus snow coverage may be more sensitive to insolation in autumn than during the rest of the year. If so, the autumn perihelion would delay the seasonal expansion of snow in the Northern Hemisphere but would have no measurable effect on the Southern Hemisphere, where seasonal snow fields are negligible.

### The astronomical climate index

As shown in Fig. 2, high obliquity, high eccentricity, and autumn perihelion correlate with relatively warm climates. Based on this observation, we formulate an index which would predict a gross climate state as a function of  $\epsilon$ ,  $\omega$  and  $e$ . The resulting astronomical climate index (ACLIN) should be high when  $\epsilon$  and  $e$  are high and when  $\omega$  approaches  $360^\circ$ . This approach leads to a purely empirical formula, ACLIN 1 ( $\alpha_1$ ), first introduced in 1978 (ref. 65):

$$\alpha_1 = \left| \frac{\omega_1 - 180}{90} \right| + \epsilon_2 - 22 + \frac{10^3 e_1^2}{2} \quad (1)$$

where  $\alpha_1$  is the climate index at time  $t_1$ , and  $\omega_1$ ,  $e_1$  the longitude of the perihelion, and eccentricity at a time  $t_1$ , and  $\epsilon_2$  the obliquity at a time  $t_2$ , defined by:

$$\omega_2 = \omega_1 - 90 \text{ and } |t_1 - t_2| = \min$$

and where  $\omega_2$  is the longitude of the perihelion at a time  $t_2$ . The separation of  $t_1$  and  $t_2$  by  $90^\circ$  in the longitude of the perihelion corresponds to a time lag of  $\sim 5,000$ – $6,000$  yr implying a several millenia long memory in the climate system. Plots of  $\alpha_1$  are compared with palaeoclimatic evidence from the past 130,000 yr in Fig. 3, and for the past and future 1 Myr in Fig. 4.

While the correlation of ACLIN 1 with palaeoclimatic evidence is good, the separation of  $e$  and  $\omega$  in two independent

variables disagrees with the laws of orbital geometry. An improved ACLIN formula therefore takes into account the physical link between orbital perturbation and insolation. It is composed of two terms:  $A$ , proportional to the relative intensity of ground level insolation at high latitudes at the summer solstice, and  $B$ , proportional to the relative intensity of insolation at the autumnal equinox.

$$A = \sin \epsilon \frac{1 + 2e \cos(\omega + 90)}{(1 - e^2)^2} \quad (2)$$

$$B = \frac{1 + 2e \cos \omega}{(1 - e^2)^2} \quad (3)$$

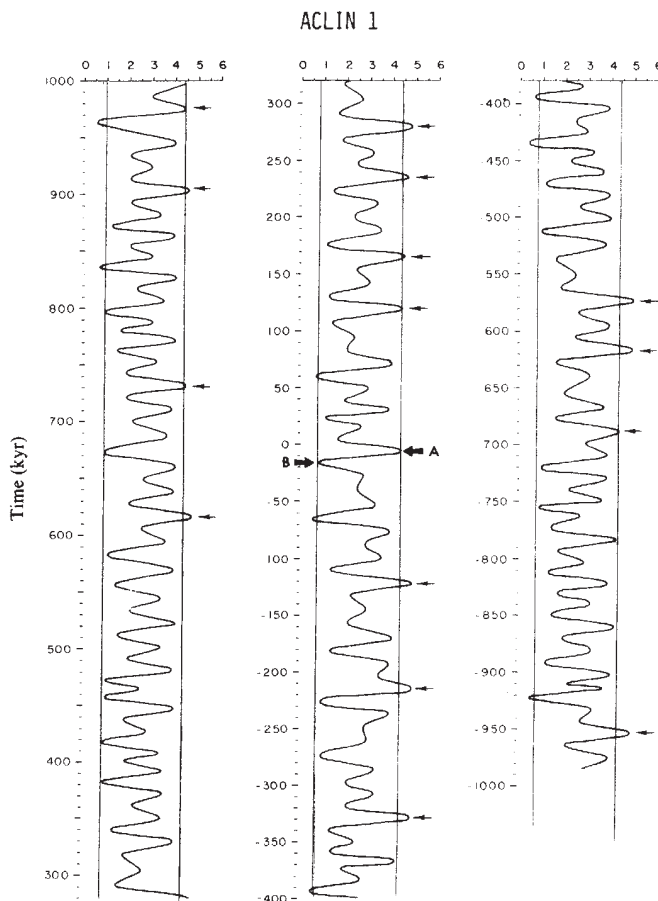
Both  $A$  and  $B$  are scanned at the same times  $t_1$  and  $t_2$  as in ACLIN 1. Other separations of  $t_1$  and  $t_2$  are possible but have not been investigated (see ref. 6). The present arbitrary selection of  $t_2$  as used in all ACLIN formulas was based on the good correlation of EOS with the climate index. As the orbital variables are scanned at two different intervals, ACLIN is a time-lag bivariate model.

The basic ACLIN formula has the simplified form

$$\alpha_0 = a(A_1 + b \cdot B_1) + (A_2 + c \cdot B_2) \quad (4)$$

and can be written in full as

$$\alpha_0 = a \frac{\sin \epsilon_1 [1 + 2e_1 \cos(\omega_1 + 90)] + b(1 + 2e_1 \cos \omega_1)}{(1 - e_1^2)^2} + \frac{\sin \epsilon_2 [1 + 2e_2 \cos(\omega_2 + 90)] + c(1 + 2e_2 \cos \omega_2)}{(1 - e_2^2)^2} \quad (5)$$



**Fig. 4** Plot of ACLIN 1 from 1,000,000 yr BP to 1,000,000 yr AP. Higher value indicates warmer climate. Peak Holocene (A) level of  $\geq 4.3$  defines interglacials marked with arrows. Past peak glacial level at about 17,000 yr BP (B).



where  $\alpha_0$  is the ACLIN index for time  $t_1$ ,  $e_1, \varepsilon_1, \omega_1$  the orbital elements at a time  $t_1$  and  $e_2, \varepsilon_2, \omega_2$  the orbital elements at a preceding interval  $t_2$  defined in the same way as in ACLIN 1, and  $a, b$  and  $c$  the arbitrary coefficients.

Various versions of  $\alpha_0$  were studied. Best results were obtained with ACLIN 2 ( $\alpha_2$ ), in which  $a = b = 0.25$  and  $c = a^2$ . A plot of ACLIN 2 is shown in Fig. 3.

The following criteria have been applied to test the predictive merits of ACLIN variants:

Accuracy level 1: predicts 10 interglacials in the past Myr, spread at  $\sim 10^5$  yr intervals, it also fulfills criteria for level 3.

Accuracy level 2: predicts interglacials in the past 350,000 yr at times matching palaeoclimatic records within the accuracy of radiometric dating. Predicts temperate interstadials (and not interglacials) at  $\sim 80,000$  and  $100,000$  yr BP and the relative amplitudes of cold minima at  $\sim 18,000, 65,000$  and  $110,000$  yr BP.

Accuracy level 3: predicts the relative amplitude and timing of interglacials and temperate interstadials during the past 150,000 yr and the timing but not the relative amplitude of the cold peaks.

Accuracy level 4: predicts correctly the relative timing but not the amplitude of warm and cold intervals during the past 150,000 yr. Does not predict interglacials.

No solution has yet been found reaching accuracy level 1. ACLIN 1 is the only formula tested so far which reached accuracy level 2. Formula ACLIN 2 reached accuracy level 3. Additional ACLIN variants can be constructed to obtain a better agreement with palaeoclimatic records, following the tuning strategy proposed by Imbrie and Imbrie<sup>8</sup>. However, the two described versions are sufficiently successful to be used for long range climate estimates.

ACLIN time series approach a smooth sinusoidal function. A computer program can transform ACLIN into a stepwise form, scanning for maxima and minima (Table 2), and assigning a parametrized duration to the three mild climate states. In such a model, protostadials and stadials occupy the rest of the time. Protostadials are gross cold climate episodes characterized by a restricted size of continental ice sheets which occur within the first 50,000 yr or so of an interglacial-glacial cycle.

Figure 5 shows the spectral response of ACLIN 1 and ACLIN 2 compared with amplitude spectra of the deep sea core V28-238 and the spectrum of Imbrie and Imbrie's model. ACLIN 2 has significant peaks only at periods of  $\sim 41,000$  yr (the obliquity cycle) and  $\sim 23,000$  and  $19,000$  yr (the precession cycle) whereas ACLIN 1 shows an additional significant peak near  $100,000$  yr. Although the  $100,000$ -yr peak in ACLIN 1 is not as dominant as it is in the oxygen isotope record of core V28-238, we find the ACLIN 1 result encouraging. It shows that a bivariate model based solely on the orbital input is capable of generating a  $100,000$ -yr cycle. In ACLIN this period is due to the eccentricity, but the real cause is the coupling of high obliquity and precession modulated by the eccentricity (see Table 1).

The spectrum of ACLIN 1, based entirely on orbital variables, closely resembles the spectrum of Imbrie and Imbrie's model<sup>8</sup> in which nonlinear ice growth rates were combined with an insolation function. Similar to Imbrie's model, ACLIN 1 has a power in the  $400,000$ -yr range but with a lower and more realistic peak.

## Discussion

Given the complexity of the Earth-atmosphere system, it is surprising that the simple ACLIN formulae successfully simulate the principal features of the palaeoclimatic record. This emphasizes again the dominant role of orbital variations in the shaping of past climates. The most significant achievement of ACLIN is the successful discrimination of interglacials on the sole basis of orbital input scanned across a  $\sim 5,000$ -yr interval. All earlier models had to include a nonlinear response of the terrestrial system to orbital forcing to separate the interglacials and generate the  $100,000$ -yr periodicity.

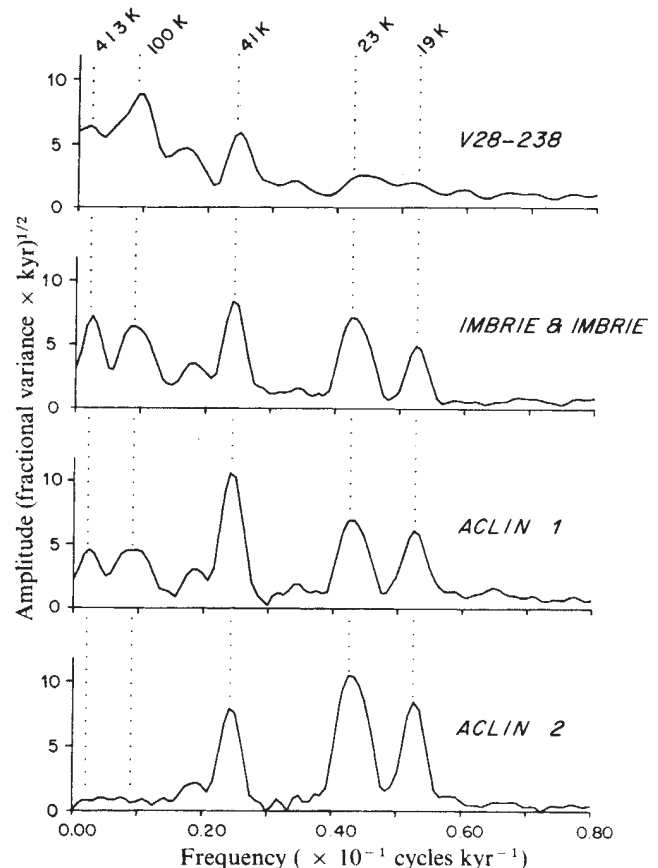


Fig. 5 Power spectra of the oxygen isotope record in the Pacific core V28-238 over the past 730,000 yr (ref. 59); the output of the Imbrie and Imbrie non-linear climate model<sup>8</sup> for the past 1,000,000 yr; the ACLIN 1 model over the past 1,000,000 yr; and the ACLIN 2 model over the same interval. Observe a  $\sim 100,000$  and  $\sim 410,000$  yr peak present in the ACLIN 1 and in the Imbrie model, but missing in ACLIN 2.

While utmost caution is needed in the interpretation of our data<sup>8</sup>, the results have some implications for climate theory:

- (1) The gross climate state may almost be in equilibrium with orbital forcing. Otherwise, it would seem impossible to predict interglacials solely from the orbital configuration of a  $\sim 5,000$  yr long interval.
- (2) The primary origin of the  $100,000$  yr cycle can be orbital rather than terrestrial. If this is so, a primary orbital forcing would be modulated by a nonlinear thermal, mass, and isostatic response of ice and oceans, rather than the reverse.
- (3) Increased insolation to high latitudes at the expense of low latitudes is associated with warm global climates, not the other way around.
- (4) The interaction of both hemispheres, not the sole dominance of the Northern Hemisphere, seems to be critical in the dynamics of the  $100,000$  yr cycle. This is indicated by the dominant role of obliquity in the ACLIN formula. (It had been suggested earlier<sup>1,2</sup> that the tilt may have played a principal role in shaping the past climates.)
- (5) The seasonal pack-ice in both hemispheres and the snow fields in the Northern Hemisphere seem to play a key part in the dynamics of Pleistocene climates as suggested earlier<sup>62,66</sup>.

While the ACLIN simulates the relative intensity of mild events quite well, it does not reproduce the cold episodes very well. For instance, the  $133,000$ -yr BP ACLIN 1 value is considerably higher than that for  $17,000$  yr BP, while the extent and the volume of continental ice was approximately the same at both intervals. Interglacials indicated by proxy data at  $\sim 400,000$  and  $500,000$  yr BP, are also poorly reproduced by ACLIN. ACLIN peaks at corresponding times do not reach the

4.3 cutoff level. However, the accuracy of both geological and astronomical data before 350,000 yr BP is low.

The general agreement of the ACLIN time series with the reconstructed palaeoclimates permits future gross climate states to be predicted. Such an estimate, however, can only be valid if man's impact on the land and the atmosphere has not significantly modified the mechanism of climate change, and will not do so in the future.

The ACLIN time series is extrapolated into the future in Figs 3 and 4. Both  $\alpha_1$  and  $\alpha_2$  predict a considerably cooler climate for the next 114,000 yr, followed by an interglacial. The first cold peak with a period of 10,000–100,000 yr should occur 4,000 yr AP (after present). ACLIN 1, but not ACLIN 2, predicts an amelioration around 15,000 yr AP followed by a cold interval centred around 23,000 yr AP. This agrees with the

prediction of Berger *et al.*<sup>67</sup> and Imbrie and Imbrie<sup>8</sup>. A major glaciation, comparable to stage 4 of the last glacial cycle, is indicated at 60,000 yr AP.

Our approach can predict at best only a very approximate general framework of gross climate changes. Detailed climate reconstructions or predictions will have to consider interactions of orbital perturbations with changing solar activity and with natural and manmade variables of terrestrial origin.

We thank J. Imbrie for innumerable improvements as well as for Fig. 5 data, and N. Shackleton, K. Bergman, J. Hays, and W. S. Broecker for reading the manuscript. Research was supported by NSF grants ATM77-28522 and ATM80-01470, and the Office of Nuclear Waste Isolation, managed by Battelle Memorial Institute for the Department of Energy under contract DE-AC06-76-RLO-1830.

Received 27 May; accepted 8 December 1980.

1. Emiliani, G. & Geiss, J. *Geol. Rdsch.* **46**, 576–601 (1959).
2. Broecker, W. S. *Science* **151**, 299–304 (1966).
3. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
4. Berger, A. *Nature* **269**, 44–45 (1977).
5. Goreau, T. J. *Nature* **287**, 620–622 (1980).
6. Calder, N. *Nature* **252**, 216–218 (1974).
7. Chin, W. Q. & Yevjevich, V. *Hydrology Papers* 65 (Colorado State University, 1974).
8. Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
9. Oerlemans, J. *Nature* **287**, 436–438 (1980).
10. Mitchell, J. M. Jr in *The Quaternary of the United States, INQUA VII Cong.* (eds Wright, H. E. Jr & Frey, D. G.) 381–901 (Princeton University Press, 1965).
11. Flint, R. F. *Glacial and Quaternary Geology* (Wiley, New York, 1971).
12. Woldstedt, P. *Das Eiszeitalter* 1, 2, 3 (Enke, Stuttgart, 1954).
13. Emiliani, C. J. *Geol.* **74**, 109–124 (1966).
14. Ruddiman, W. F. & McIntyre, A. *Geol. Soc. Am. Mem.* **145**, 111–146 (1976).
15. Kukla, G. *Earth Sci. Rev.* **13**, 307–374 (1977).
16. Grootes, P. M. *Science* **200**, 11–15 (1978).
17. Woillard, G. in *IGCP Prof. 73/1/24, Rep. 6* (ed. Sibrava, V.) (CSSR, 1980).
18. Woillard, G. *Acta geogr. slov.* **14**, 1–118 (1975).
19. Woillard, G. *Quat. Res.* **9**, 1–21 (1978).
20. Woillard, G. *Bull. Soc. belge Géol.* **88**, 51–69 (1979).
21. Wijmstra, T. A. *Acta bot. neerl.* **4**, 511–527 (1969).
22. Wijmstra, T. A., & Smit, A. *Acta bot. neerl.* **25**, 297–312 (1976).
23. Broecker, W. S. *et al. Science* **159**, 297–300 (1968).
24. Harmon, R. S., Ku, T.-L., Matthews, R. K. & Smart, P. L. *Geology* **7**, 405–409 (1979).
25. James, N., Mountjoy, E. & Omura, A. *Bull. geol. Soc. Am.* **82**, 2011–2018 (1971).
26. Steinen, R. P., Harrison, R. S. & Matthews, R. K. *Bull. geol. Soc. Am.* **84**, 63–70 (1973).
27. Mesolella, K. T., Matthews, R. K., Broecker, W. S. & Thurber, D. L. *J. Geol.* **77**, 250–274 (1969).
28. Fairbanks, R. K. & Matthews, R. K. *Quat. Res.* **10**, 181–196 (1978).
29. Matthews, R. K. *Quat. Res.* **3**, 147–153 (1973).
30. Bender, M. L. *et al. Bull. geol. Soc. Am. Part I*, **90**, 577–594 (1979).
31. Land, L. S., MacKenzie, F. T. & Gould, S. J. *Bull. geol. Soc. Am.* **78**, 993–1006 (1967).
32. Harmon, R. S., Schwarcz, H. P. & Ford, D. C. *Quat. Res.* **9**, 205–218 (1978).
33. Veeh, H. H. & Chappell, J. M. A. *Science* **167**, 862–865 (1970).
34. Chappell, J. & Veeh, H. H. *Nature* **276**, 602–604 (1978).
35. Chappell, J. *Bull. geol. Soc. Am.* **85**, 553–570 (1974).
36. Bloom, A. L., Broecker, W. S., Chappell, J. M. A., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 185–205 (1974).
37. Konishi, K., Schlanger, S. & Omura, A. *Mar. Geol.* **9**, 225–240 (1970).
38. Marshall, J. F. & Launay, J. *Quat. Res.* **9**, 186–192 (1978).
39. Veeh, H. H., Schwabel, D., Van de Graaf, W. J. E. & Denman, P. D. J. *geol. Soc. Austr.* **26**, 285–292 (1979).
40. Thurber, D., Broecker, W., Blanchard, R. & Potratz, H. *Science* **149**, 55–58 (1965).
41. Broecker, W. & Thurber, D. *Science* **149**, 58–60 (1965).
42. Neumann, A. C., & Moore, W. S. *Quat. Res.* **5**, 215–224 (1975).
43. Emery, K. O. & Merrill, A. S. *Bull. geol. Soc. Am.* **90**, 689–694 (1979).
44. MacIntyre, I. G., Pilkey, O. H. & Stuckenrath, R. *Bull. geol. Soc. Am.* **90**, 692–694 (1979).
45. Cullingford, R. A., Caseldine, C. J. & Gotts, P. E. *Nature* **284**, 159–161 (1980).
46. Jones, B. G., Young, R. W., & Eliot, I. G. *J. geol. Soc. Austr.* **26**, 255–264 (1979).
47. Shackleton, N. J. in *The Fate of Fossil Fuel* (eds Anderson, N. & Malahoff, A.) 401–427 (Plenum, New York, 1977).
48. Ninkovich, D. & Shackleton, N. J. *Earth planet. Sci. Lett.* **27**, 20–34 (1975).
49. Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168–171 (1979).
50. Ninkovich, D., Shackleton, N. J., Abdel-Monem, A. A., Obradovich, J. D. & Izett, G. *Nature* **276**, 574–577 (1978).
51. Van Der Hammen, T., Wijmstra, T. A. & Zagwijn, W. H. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 391–424 (Yale University Press, New Haven, 1971).
52. Woillard, G. *Nature* **281**, 558–562 (1979).
53. Bryson, R. A. & Wendland, W. M. in *Life, Land and Water* (ed. Mayer-Oakes, W. J.) 271–298 (University of Manitoba Press, Winnipeg, 1967).
54. Broecker, W. S., Ewing, M. & Heezen, B. C. *Am. J. Sci.* **258**, 429–448 (1960).
55. Morner, N.-A. *Can. J. Earth Sci.* **8**, 1423–1431 (1971).
56. Kukla, G. in *Paleoecology of Africa* (ed. van Zinderen, E. M.) (Bakker, The Netherlands, in the press).
57. Morner, N.-A. & Dreimanis, A. *Geol. Soc. Am. Mem.* **136**, 107–134 (1973).
58. Dreimanis, A. & Goldthwait, R. P. *Geol. Soc. Am. Mem.* **136**, 71–106 (1973).
59. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39 (1973).
60. Berger, A. *Astr. Astrophys.* **51**, 127–135 (1976).
61. Berger, A. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
62. Kukla, G. J. in *Climate Change* (ed. Gribbin, J.) 114–129. (Cambridge University Press, 1978).
63. Washington, W. M. & Chervin, R. M. *J. Appl. Met.* **18**, 3–16 (1979).
64. Kukla, G. J. *Nature* **253**, 600–603 (1975).
65. Kukla, G. in *Probability of expected climate stresses in North America in the next one million years: Topical Report* (Battelle, Pacific Northwest Laboratories, 1978).
66. Croll, J. *Climate and Time in the Geological Relations: A Theory of Secular Changes of the Earth's Climate* 4th edn (London, 1975).
67. Berger, A., Guioit, J. & Kukla, G. in *Proc. int. Scientific Assembly, Serbian Acad. Sci. Arts, 1979* (in the press).
68. Milankovitch, M. M. *Beograd, Koninglich Serbische Adademie* 484 (English transl. by Israel Program for Scientific Translation and published for the U.S. Department of Commerce and the National Science Foundation, Washington DC, 1941).

# A carbon budget for overfishing off Peru

John J. Walsh

Oceanographic Sciences Division, Brookhaven National Laboratory, Upton, New York 11973, USA

*The rise and collapse of the Peruvian anchovy fishery between the 1950s and the 1970s was accompanied by marked changes in the fluxes of a carbon budget for the upwelling ecosystem.*

THE anchovy harvest off the coast of Peru increased from 700,000 tonnes  $\text{yr}^{-1}$  in 1958 to a maximum of  $\sim 12$  million tonnes  $\text{yr}^{-1}$  in 1970, but then declined precipitously to  $\sim 2$  million tonnes  $\text{yr}^{-1}$  in 1973 and to  $\sim 1$  million tonnes  $\text{yr}^{-1}$  between 1977 and 1979. The collapse of the anchovy fishery, located in the upwelling ecosystem off the coast of Peru, is attributable to overfishing and the diminished survival of anchovy larvae during warm-water intrusions (El Niños) southwards along the coast of Peru in both 1972 and 1976<sup>1</sup>.

The very large yield of anchovy (*Engraulis ringens*) in the heyday of the fishery is usually attributed to high primary production in the upwelling system and the high efficiency of the

food chain in which the commercial fish feed directly on phytoplankton. In reality, there has been debate<sup>2,3</sup> for more than a decade about the relative importance in the maintenance of the anchovy stocks of the short food web (phytoplankton  $\rightarrow$  anchovy) with  $\sim 10\%$  ecological efficiency and of a longer food web (phytoplankton  $\rightarrow$  zooplankton  $\rightarrow$  anchovy) with efficiency of  $\sim 1\%$  but functioning over an area perhaps 10 times as great. A primitive simulation<sup>4</sup> of the daily nitrogen flux through the Peruvian upwelling system suggested that the longer and more widely spread food-web might be more important during the strong winter upwelling off Peru. A clearer understanding of the trophodynamics of the upwelling system is now possible by



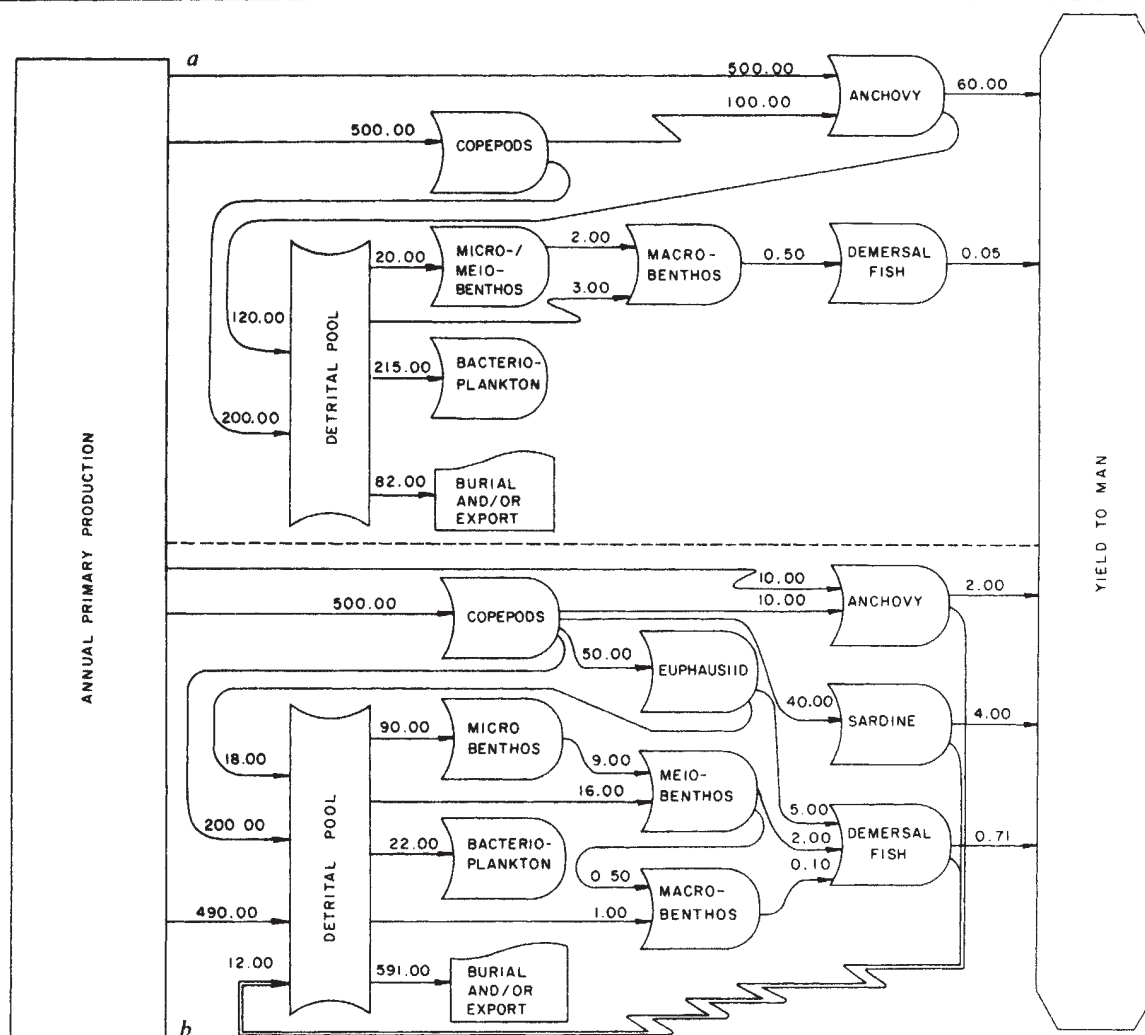


Fig. 1 A budget for the flux of carbon ( $\text{g C m}^{-2} \text{yr}^{-1}$ ) through the Peru food web: a, before (~1966-69) and b, after (~1976-79) overfishing of anchovy.

examining the annual carbon budgets before and after overfishing, in the periods 1966-69 and 1976-79 respectively. The 1966 primary production data (Table 1) are high because they were collected while the ship was drifting within a high chlorophyll area at the shelf-break<sup>4</sup>. The March-April 1976-77 data are low because repeated time series stations were taken within a low chlorophyll area near the coast<sup>1</sup>. Other data in Table 1 were averaged from transects taken across the shelf. High chlorophyll and abundant anchovy eggs are found within 50 km of the coast<sup>1</sup>, suggesting that a combined productivity average of this region is a reasonable estimate of the yearly carbon input to the Peru upwelling food chain. At a rate of at least  $3 \text{ g C m}^{-2} \text{day}^{-1}$  ( $\bar{X} = 4 \text{ g C m}^{-2} \text{day}^{-1}$  in Table 1), the annual primary production is a minimum of  $\sim 1,000 \text{ g C m}^{-2} \text{yr}^{-1}$  off Peru (see Table 2) in contrast to  $\sim 200\text{--}500 \text{ g C m}^{-2} \text{yr}^{-1}$  for shelf ecosystems at higher latitudes<sup>5</sup>.

An upwelling zone of  $\sim 50 \text{ km}$  offshore width, extending along the Peru-Chile coast of  $\sim 2,000 \text{ km}$  length ( $4\text{--}22^\circ\text{S}$ ), and a peak fish catch of 10 million tonnes  $\text{yr}^{-1}$  harvested over the whole area ( $1 \times 10^5 \text{ km}^2$ ) suggests a yield of at least 100 tonnes  $\text{km}^{-2} \text{yr}^{-1}$ . Such a harvest was an order of magnitude higher than that presently derived from the North Sea, the Mid-Atlantic Bight or the Bering Sea<sup>5</sup>. With a conversion factor of  $0.06 \text{ g C} = 1 \text{ g ww}$ , a harvest of 100 tonnes  $\text{km}^{-2} \text{yr}^{-1}$  ( $100 \text{ g m}^{-2} \text{yr}^{-1}$ ) suggests, however, a yield to man of only  $6 \text{ g C m}^{-2} \text{yr}^{-1}$  of anchovy off Peru from a primary production of at least  $\sim 1,000 \text{ g C m}^{-2} \text{yr}^{-1}$ . As man was the major consumer of the anchovy during its peak and declining yields, this harvest may be a reasonable estimate of fish secondary production over the

whole region. It suggests a food chain efficiency of  $< 1\%$  before the fishery collapsed. The yield from the anchovy fishery in 1976-79 was  $\sim 1$  million tonnes  $\text{yr}^{-1}$ , or an implied food chain efficiency to man of only  $0.1\%$  after overfishing if the catch then also approximated secondary production of the anchovy.

The same maximum estimate of secondary production, taken from a 90% smaller upwelling region, would represent a larger yield to man per unit area—a greater implied food chain efficiency. For example, if most of the peak yield had been derived only from the area of high anchovy egg and larvae abundance<sup>1</sup>, north of Callao on the broad shelf between  $6$  and  $10^\circ\text{S}$ , then the annual harvest would have been  $60 \text{ g C m}^{-2} \text{yr}^{-1}$  of fish before overfishing occurred (Table 2). If the anchovy in this area also consumed phytoplankton half the year and zooplankton with a growth efficiency of  $20\%$  (Table 2) during the other half of the year<sup>6</sup>, the fish yield would then have been  $10\%$  of their food input from this smaller region between  $6$  and  $10^\circ\text{S}$ . A more detailed carbon budget for the food web off Peru may determine the fate of the primary production along  $90\%$  of the Peru coast before overfishing, if most of the annual yield was derived from facultative herbivorous fish within an area of only  $1 \times 10^4 \text{ km}^2$ .

### Carbon budget off Peru

Information on oxygen content, particle fluxes to the bottom, sea-bed respiration, organic content of the sediments and nutrient depletion of the water column aid in the construction of a more complete carbon budget before (Fig. 1a) and after overfishing (Fig. 1b). Previous analysis of the demersal part of

the nearshore food chain before overfishing suggested that the input of carbon to the bottom consisted mainly of faecal pellets rather than phytodetritus<sup>4</sup>. In terms of carbon consumption, the decomposition of the faecal pellets could be attributed to caprophagy and/or bacterial activity in both the water column and sediments. Additional calculations<sup>7</sup> had indicated that the oxygen demand of this detrital carbon flux was such that most of the faecal pellets could be remineralized within the upper 60 m of the water column. After overfishing and the decline of grazing pressure from adult anchovy, however, a large flux of phyto-detritus may have sunk out of the water column when nutrient depletion occurred at the edge of the shelf.

The estimate for bacterioplankton production (Table 2) from consumption of faecal pellets within the upper part of the water column before overfishing (Fig. 1a) was obtained from a steady-state budget of the vertical decline of oxygen at 15°S in 1966–69; 4.5 ml O<sub>2</sub> l<sup>-1</sup> above 20 m and 0.5 ml O<sub>2</sub> l<sup>-1</sup> below<sup>7</sup>. As a result of oxygen depletion in upwelled source water during March to April 1976, 1977 and 1978, however, the average O<sub>2</sub> concentration at 15°S was ~1 ml O<sub>2</sub> l<sup>-1</sup> in the upper 20 and 0.5 ml O<sub>2</sub> l<sup>-1</sup> below, that is a vertical gradient of ~0.04 ml O<sub>2</sub> l<sup>-1</sup> m<sup>-1</sup> after overfishing in contrast to 0.4 ml O<sub>2</sub> l<sup>-1</sup> m<sup>-1</sup> a decade earlier. The local production of aerobic bacteria was thus estimated to be an order of magnitude less (Table 2) in the low chlorophyllic surface waters of the second carbon budget (Fig. 1b) with the remaining detrital carbon then sinking to where more denitrification<sup>8</sup> occurred in deeper water and sulphate reduction<sup>9</sup> in the bottom sediments.

**Table 1** Mean daily primary production in Peru Shelf waters from 1966 to 1978

| Year | g C m <sup>-2</sup> day <sup>-1</sup> | (N)  | Month            | Area    | Ref. |
|------|---------------------------------------|------|------------------|---------|------|
| 1966 | 6.3                                   | (15) | March–April      | 15°S    | 35   |
| 1969 | 5.2                                   | (32) | March–April      | 15°S    | 35   |
| 1969 | 1.3                                   | (12) | June             | 10–15°S | 4    |
| 1974 | 4.5                                   | (5)  | February–March   | 7–8°S   | 19   |
| 1975 | 3.8                                   | (5)  | March–May        | 6–12°S  | 36   |
| 1976 | 3.2                                   | (4)  | August–September | 10–15°S | 1    |
| 1976 | 1.7                                   | (29) | March–April      | 15°S    | 35   |
| 1977 | 1.9                                   | (43) | March–April      | 15°S    | 35   |
| 1977 | 7.3                                   | (10) | November         | 9–10°S  | 18   |
| 1978 | 4.3                                   | (30) | March–April      | 15°S    | 35   |

$$\bar{X} = 3.95.$$

The aerobic carbon production of the benthos before overfishing (Fig. 1a) was estimated from both P/B estimates and bottom respiration measurements at 15°S in March to April 1969<sup>10</sup>. The mean macrobenthos biomass across the shelf in 1969<sup>11</sup> was 0.5 g C m<sup>-2</sup>, suggesting a secondary production of 0.5 g C m<sup>-2</sup> yr<sup>-1</sup> (Table 2) and an ingestion demand of 5.0 g C m<sup>-2</sup> yr<sup>-1</sup> from a P/B ratio of 1 and a growth efficiency of 10%. Approximately 20% of the measured ~25 g C m<sup>-2</sup> yr<sup>-1</sup> bottom respiration<sup>10</sup> might thus be due to macrobenthos off Peru in 1969. The rest is attributed to micro- and meio-benthos as in the bottom metabolism of other shelves<sup>12,13</sup>.

The total input of carbon to the shelf bottom before overfishing (Table 2) is similar to a flux of 88 g C m<sup>-2</sup> yr<sup>-1</sup> at 50 m measured with floating sediment traps on the narrow shelf at 15°S in 1977<sup>9</sup>. If a higher sinking loss was applied (without other losses) to an early model of the nearshore Peru upwelling system<sup>14</sup>, the predicted amount of downstream chlorophyll in the offshore region was less than that observed. The nearshore area at 15°S contained the same amount of low chlorophyll (<2 µg l<sup>-1</sup>) during 1966, 1969, 1976, 1977 and 1978, as that found inshore at 10°S during 1960–69 and 1976–77, suggesting that an estimate of a sinking loss of only 10% of nearshore primary production is reasonable, as higher phytoplankton biomass is found downstream, that is, offshore.

The production of the benthic food web after overfishing (Fig. 1b) was first estimated from sulphate reduction estimates of 90 g C m<sup>-2</sup> yr<sup>-1</sup> for anaerobic metabolic demand at 15°S in

surface sediments<sup>9</sup>. The sulphur bacterium, *Thioploca*, was found on the bottom in large numbers along the whole Peru coast during 1976, and the near bottom oxygen concentration at 15°S on the 100-m isobath was <0.1 ml O<sub>2</sub> l<sup>-1</sup> in 1976–77, compared with ~0.2–1.0 ml O<sub>2</sub> l<sup>-1</sup> in 1969. The content of the secondary nitrite maximum increased from 3–6 µg-atom NO<sub>2</sub> l<sup>-1</sup> in 1966–69<sup>15</sup> to 14 µg-atom NO<sub>2</sub> l<sup>-1</sup> in 1977<sup>8</sup>. Similarly no sulphate reduction was detected in the water column in 1969<sup>16</sup> at 15°S, whereas H<sub>2</sub>S was found there in March to April 1976<sup>17</sup> and at 8–9°S during November 1977<sup>18</sup>. Based on an average 1977 meio-benthic biomass of 0.6 g C m<sup>-2</sup> across the shelf<sup>9</sup> and a P/B ratio of 4/1 reflecting a higher turnover rate of these small organisms, the secondary production of the meio-benthos might have been 2.5 g C m<sup>-2</sup> yr<sup>-1</sup> (Table 2), 28% of the sulphate reduction estimate of microbial production. Quantitative macrobenthic data are not available for 1976–78 but their biomass and production (Table 2) were then presumably smaller because more extensive anoxic conditions occurred on the Peru shelf.

As the primary electron acceptor in the oxidation of organic matter in the sea is oxygen, then nitrate in the process of denitrification, and finally sulphate reduction, the above sequence suggests that a larger than normal organic loading to the bottom of the shelf and upper slope had occurred by 1976–77 (Fig. 1b). The total detrital flux from the water column after overfishing in 1976–79 is, in fact, estimated (Table 2) to be 70% of the 1,000 g C m<sup>-2</sup> yr<sup>-1</sup> primary production. After the 1976 El Niño, such a detrital loss (Fig. 1b) is almost an order of magnitude larger than that estimated for 1966–69 (Fig. 1a), but similar to an estimate of 64% detrital loss<sup>19</sup> in 1974 at 7–8°S after the 1972 El Niño. Within the geological record, a 10-fold increase in sedimentation rate is associated with a doubling of the organic carbon content of the sediments<sup>20</sup>. If the above carbon budgets are correct, a similar process may have occurred off Peru from 1966–69 to 1976–79.

The distribution of organic carbon in the surface sediments (Fig. 2) suggests that before overfishing in 1968 and 1969<sup>11,21,22</sup> the sediment carbon off Peru increased both seawards at the slope and polewards. Without consideration of changes in food chain dynamics, a primary production rate of 1,000 g C m<sup>-2</sup> yr<sup>-1</sup> and a previous sedimentation rate of 0.1 cm yr<sup>-1</sup> can be calculated<sup>20</sup> as a carbon accumulation rate within the sediments of 114 g C m<sup>-2</sup> yr<sup>-1</sup> off Peru, that is similar to that from the carbon budget before overfishing (Table 2). The location of the 2% carbon isopleth in the sediments of the wide shelf north of Callao in 1968–69 (Fig. 2) also suggests that less carbon used to be lost from the water column at 10°S than 15°S. These spatial trends continued after overfishing in 1976 (C. Delgado, L. A. Flores and G. T. Rowe, unpublished data) and 1977<sup>18,23</sup>, but the carbon content of the surface sediments may have doubled recently. The same primary production and a sixfold increase in sedimentation rate (Table 2) lead to an empirical carbon accumulation rate<sup>20</sup> of 684 g C m<sup>-2</sup> yr<sup>-1</sup> in the sediments compared with that of 591 g C m<sup>-2</sup> yr<sup>-1</sup> estimated from the overfishing budget (Fig. 1b).

At 15°S, 4.5% carbon was the highest surface sediment value observed on the shelf in March to April 1969<sup>11</sup>, but 10% carbon was found at the same stations with the same methods in 1976. Furthermore, within 22 cores taken across the 15°S Peru shelf and slope in 1977–78, the mean subsurface C/N value at 50 cm depth of these cores was 7.76 compared with 7.53 in the surface sediments, suggesting little diagenetic change<sup>24</sup>. However, the mean % carbon was 2.89 at 50 cm depth and 4.82 at the surface, implying a recent increase in carbon flux to the sediments that could have occurred within the past decade. Inadequate spatial sampling means that areas of high organic content may have been missed in these 1968–78 surveys, but the horizontal, temporal and vertical trends of the sediment carbon data are at least consistent with the carbon budgets of the water column. The first carbon budget (Fig. 1a) for the food web before overfishing may be most applicable to northern Peru waters during 1966–69. The second budget (Fig. 1b) may instead reflect



the food web both before overfishing off southern Peru and afterwards off northern Peru.

The predicted change in dominance of pelagic and demersal fish species after overfishing (Fig. 1b) reflects the decline in these budgets of anchovy ingestion demands on both phytoplankton and zooplankton<sup>6</sup>. The increase of zooplankton biomass after the 1972 and 1976 El Niños was mainly in the standing stocks of omnivorous euphausiids<sup>1,19</sup>. In contrast, there was little apparent change in their biomass after the 1957 El Niño<sup>25</sup>, when the anchovy fishery was just beginning. The sardine, *Sardinops sagax*, and hake, *Merluccius gayi*, both eat zooplankton off Peru, suggesting that the stocks of these fish should also increase after the anchovy decline. Similar changes occurred in demersal species abundance within the North Sea after the pelagic fish stocks were overfished<sup>26</sup>.

Following the 1976 El Niño, sardine indeed constituted ~67% of the pelagic yield of 1.5 million tonnes  $\text{yr}^{-1}$  in 1977. During 1978 and 1979, the catch of mackerel, sardine, pipefish, and horse-mackerel was still ~50% of the total pelagic yield of 1–2 million tonnes  $\text{yr}^{-1}$ . Furthermore, the potential demersal yield of 'virgin' hake stocks was estimated to be 0.75 million tonnes  $\text{yr}^{-1}$  after the 1972 El Niño<sup>27</sup>. This is similar to the change predicted in the second carbon budget (Fig. 1b) of a yield of 0.71  $\text{g C m}^{-2} \text{yr}^{-1}$ , or 1.2 million tonnes  $\text{yr}^{-1}$  of demersal fish along the whole coast ( $1 \times 10^5 \text{ km}^2$ ). Despite restrictions in fishing, however, the same amount of plankton was not consumed by euphausiids, sardine, mackerel and hake in 1976 as by anchovy in 1966. Instead, the oxygen content of the euphotic zone and sea bed has since declined, the nitrite and sulphide content of the water column increased, and the organic content of the bottom sediments also apparently increased after overfishing.

### Nitrogen mass balance for carbon fluxes

If  $10^3$  tonnes  $\text{km}^{-2} \text{yr}^{-1}$  of anchovy had been produced on the wide shelf north of Callao ( $1 \times 10^4 \text{ km}^2$ ) and  $10^2$  tonnes  $\text{km}^{-2} \text{yr}^{-1}$  along the rest of the coast ( $9 \times 10^4 \text{ km}^2$ ), the combined secondary production of 19 million tonnes  $\text{yr}^{-1}$  before overfishing is the same as that originally estimated in 1969<sup>2,3</sup>. The carbon budget for the Peru food chain after overfishing (Fig. 1b) is thus a more reasonable description of the food chain along most of the Peru coast before and after overfishing with the exception of the 6–10°S area. If this second budget did apply to 15°S during 1966–69 (before overfishing), then about half of the daily production of phytoplankton was always lost as sinking detritus to the upper slope after nutrient depletion within offshore waters<sup>28</sup> rather than cropped inshore by zooplankton and anchovy<sup>4,29</sup>. Annual carbon budgets of the shelf food webs of the Mid-Atlantic Bight<sup>5</sup>, the Gulf of Mexico and the Bering Sea similarly suggest that ~50% of their primary production

**Table 2** Changes in annual carbon fluxes of the Peru upwelling ecosystem between 1966–69 and 1976–79

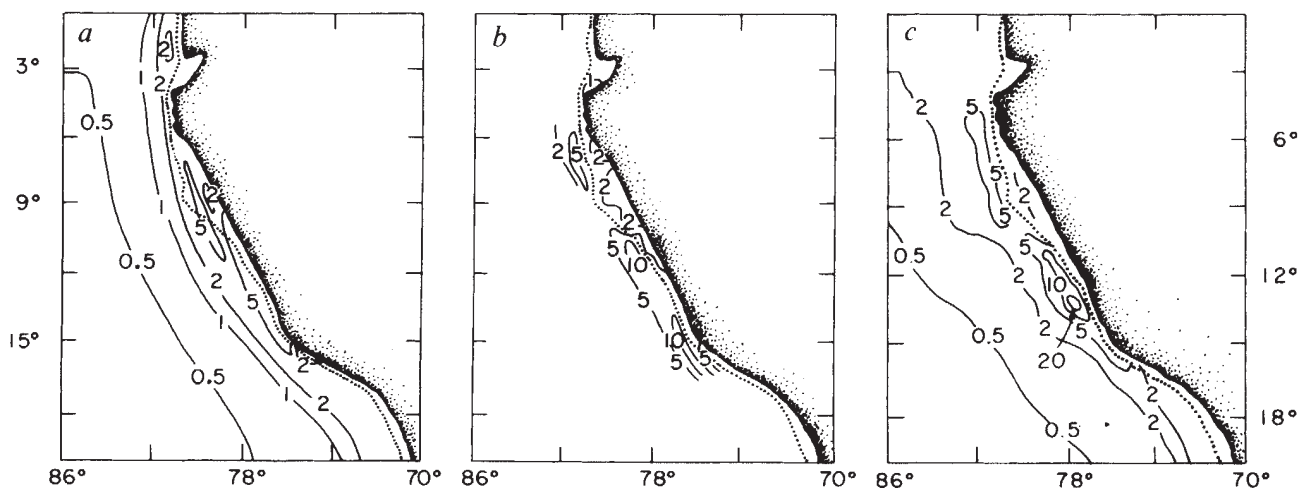
|                              | 1966–69 (Ref.)<br>( $\text{g C m}^{-2} \text{yr}^{-1}$ ) | 1976–79 (Ref.)<br>( $\text{g C m}^{-2} \text{yr}^{-1}$ ) |
|------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Primary production           | 1570* (4, 35)                                            | 1351* (1, 35, 18)                                        |
| Copepod production           | 100†                                                     | 100†                                                     |
| Euphausiid production        | 2.5†                                                     | 5†                                                       |
| Anchovy yield                | 60*†                                                     | 6*†                                                      |
| Sardine yield                | 0.4*                                                     | 4*                                                       |
| Hake yield                   | 0.05†                                                    | 0.7* (27)                                                |
| Bacterioplankton production  | 215† (7)                                                 | 22†                                                      |
| Microbenthic production      | } 2*† (10)                                               | 9*† (9)                                                  |
| Meiobenthic production       |                                                          | 2.5*† (9)                                                |
| Macroinvertebrate production | 0.5*† (11)                                               | 0.1†                                                     |
| Detrital carbon production   | 320† (4)                                                 | 720†                                                     |
| Sinking loss                 | 105†                                                     | 698†                                                     |
| Sediment storage/export      | 82†                                                      | 591†                                                     |

\* Measured; † calculated.

may be exported to the continental slope as phytodetrital carbon from these shelves, thus constituting a major organic sink of the global  $\text{CO}_2$  budget<sup>30</sup>.

An annual detrital carbon loss of ~500  $\text{g C m}^{-2} \text{yr}^{-1}$  at 15°S, and a C/N ratio of 5/1 implies that the particulate nitrogen export in the form of detritus is 100  $\text{g N m}^{-2} \text{yr}^{-1}$ , in contrast to ~1  $\text{g N m}^{-2} \text{yr}^{-1}$  in the form of fish yield (Fig. 1b). To maintain a nitrogen mass balance in an annual steady state of the food web on this Peru shelf, ~100  $\text{g N m}^{-2} \text{yr}^{-1}$  of dissolved nitrogen must at least be returned to the shelf to replace the detrital and fish nitrogen loss. Furthermore, an annual primary production of ~1,000  $\text{g C m}^{-2} \text{yr}^{-1}$  at 15°S requires 200  $\text{g N m}^{-2} \text{yr}^{-1}$ , 40% of which could be supplied by uptake of ammonium from nitrogen recycling<sup>31</sup> and 60% by physical input of nitrate from upwelling, because river runoff is negligible<sup>24</sup>. An estimate of the nitrate supply can be made with a simple budget,  $\delta \text{NO}_3 / \delta t = w \delta \text{NO}_3 / \delta z$ , using an average upwelling rate,  $w$ , of 10  $\text{m day}^{-1}$  and an observed vertical nitrate gradient of 5  $\text{mg-atom NO}_3 \text{ m}^{-3}$  over the 50-m depth interval of the nearshore upwelling region off Peru<sup>4</sup>; this daily flux sums to 5  $\text{g NO}_3 \text{ m}^{-3} \text{yr}^{-1}$ . Over a 20–30-m euphotic zone, such an input of 100–150  $\text{g NO}_3 \text{ m}^{-2} \text{yr}^{-1}$  would be sufficient to balance the 'new' nitrogen demands<sup>32</sup> of photosynthesis, fish removal and detrital loss from the shelf at 15°S.

Similar to recent changes in North Sea herring stocks<sup>26</sup>, a 90% decline of anchovy production from 60 to 6  $\text{g C m}^{-2} \text{yr}^{-1}$  off northern Peru as a result of overfishing, however, would lead to a much larger absolute change in the detrital carbon (nitrogen) flux at 10°S (a  $\Delta$  of ~500 versus ~20  $\text{g C m}^{-2} \text{yr}^{-1}$ ) than a similar percentage reduction from 6.0 to 0.6  $\text{g C m}^{-2} \text{yr}^{-1}$  at



**Fig. 2** The distribution of % organic carbon in the surface sediments along the Peru coast in a, 1968–69,  $N = 63$ ; b, 1976,  $N = 97$ ; c, 1977  $N = 56$ .

15°S. The subsurface flow of water off Peru is to the south at  $\sim 1\text{--}5\text{ cm s}^{-1}$  within the Peru–Chile undercurrent at the edge of the shelf. With these differences in area, food chain efficiency, and yield, the total poleward, or downstream, detrital load of the Peru system could have approximately doubled after overfishing of the anchovy stocks between 6 and 10°S. Thus part of the apparent bottom carbon transient after overfishing at 15°S, that is, a doubling between 1968 and 1978 could be the result of sinking detritus advected south within the Peru–Chile undercurrent. If the 1966–69 anchovy grazing pressure were reapplied to the shelf ecosystem at 6–10°S, presumably the upstream source of phytodetritus would decline, and the per cent sediment carbon at 15°S might also decline if the accumulated surficial bottom carbon were oxidized or further exported to the continental rise<sup>24,30</sup>. Based on <sup>210</sup>Pb dating, cores taken at 186–370 m depth between 11 and 13°S in June 1977 suggest a sedimentation rate of  $0.05\text{--}0.16\text{ cm yr}^{-1}$  off northern Peru<sup>33</sup>, whereas cores taken at 82–250 m depth near 15°S in February–March 1978 indicate a rate of  $1.1\text{--}1.2\text{ cm yr}^{-1}$  off Southern Peru<sup>34</sup>.

## Conclusions

Although the estimates in these carbon budgets are questionable it seems that most of both the larval survival and harvest of adult anchovy used to occur within a 10% area of the Peru shelf. This is about the same per cent area of the Bering Sea from which most stocks of the Alaska pollock are harvested<sup>5</sup>. The

hypothesis that greater changes in shelf trophodynamics have occurred within the 6–10°S area than in the 15°S region seems to reflect the impact of both overfishing by man and alteration of larval survival by nature off Peru<sup>1</sup>. The decline of anchovy grazing pressure has apparently led to increases in plankton biomass, increased stocks of sardine and hake, increased carbon loading and sulphate reduction at the 'downstream' sea bed, and a decline in oxygen and nitrate content of the water column.

Although the anchovy fishery of the Peru upwelling ecosystem was one of the world's most 'studied', 'advised' and 'managed' coastal resource, societal conflicts led to its possible demise. Along the whole coast, the fish yield may now be only 10 tonnes  $\text{km}^{-2}\text{ yr}^{-1}$ , similar to that from the Mid-Atlantic Bight, North Sea and Bering Sea despite a three- to sevenfold higher primary production off Peru. Future studies will determine whether after these changes in carbon export, the carrying capacity of this system with respect to living resources will remain at present level. It is possible that the Peru upwelling ecosystem will either return to the same level of secondary production before 1972, or maintain the present decline in food chain yield such that the greater export in the form of detrital carbon rather than of fish carbon will continue.

The data collection and reduction phase of this research was supported mainly by the NSF as part of the Coastal Upwelling Ecosystem Analysis and Process and Resources of the Bering Sea programs. Additional data analysis was supported by the DOE as part of our Atlantic Coastal Ecosystem program. Drs G. Rowe and E. Suess provided sediment carbon data.

Received 30 June 1980; accepted 6 January 1981.

- Walsh, J. J. *et al.* *Deep-Sea Res.* **27**, 1 (1980).
- Cushing, D. H. *F.A.O. Fish Biol. Tech. Pap.* **84**, 1 (1969).
- Ryther, J. H. *Science* **166**, 72 (1969).
- Walsh, J. J. *Deep-Sea Res.* **22**, 201 (1975).
- Walsh, J. J. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 159 (Academic, New York, 1980).
- De Mendiola, B. R. in *The Early Life History of Fish* (ed. Blaxter, J. H. S.) 277 (Springer, Berlin, 1974).
- Walsh, J. J. in *Estuarine Research* Vol. 1 (ed. Cronin, L. E.) 617 (Academic, New York, 1975).
- Codispoti, L. A. & Packard, T. T. *J. mar. Res.* **39**, 453 (1980).
- Rowe, G. T. in *Bioproductivity of Upwelling Ecosystems* (eds Barber, R. T. & Vinogradov, M. Y.) (Springer, Berlin, in the press).
- Pamatmat, M. *Limnol. Oceanogr.* **16**, 536 (1971).
- Rowe, G. T. *Inv. Pesq.* **35**, 127 (1971).
- Smith, K. L. *Ecology* **54**, 1065 (1978).
- Smith, K. L., Rowe, G. T. & Nichols, J. A. *Estuar. Coast. mar. Sci.* **1**, 65 (1973).
- Walsh, J. J. & Dugdale, R. C. *Inv. Pesq.* **35**, 309 (1971).
- Fiadlerio, M. & Strickland, J. D. H. *Deep-Sea Res.* **26**, 187 (1965).
- Goering, J. J. & Pamatmat, M. *Inv. Pesq.* **35**, 233 (1971).
- Dugdale, R. C., Goering, J. J., Barber, R. T., Smith, R. L. & Packard, T. T. *Deep-Sea Res.* **24**, 601 (1977).
- Project ICANE* (L.A.E. DOE Dept. Ser. BI-R-78-6, Bedford Institute of Oceanography, Halifax, 1978).
- Shushkina, E. A., Vinogradov, M. Y., Sorokin, Y. I., Lebedeva, L. P. & Mikheyev, V. N. *Oceanology* **18**, 579 (1978).
- Muller, P. J. & Suess, E. *Deep-Sea Res.* **26**, 1347 (1979).
- Longvineko, N. V. & Romankevich, E. A. *Litol. Pol. Isk.* **1**, 3 (1973).
- Rodionova, K. F., Romankevich, E. A., Duzhikova, T. N. & Teklova, M. S. *Vnigni, TRUDY* 138, *Geochim. Sbornik* **5**, 118 (1973).
- Suess, E. & Muller, P. J. *Proc. Cent. natn. Rech. Scient.* (in the press).
- Walsh, J. J., Premuzic, E. T. & Whitledge, T. E. in *Ecohydrodynamics* (ed. Nihoul, J. C. J.) (Elsevier, Amsterdam, in the press).
- Brinton, E. *Bull. Scripps Inst. Oceanogr.* **8**, 51 (1962).
- Burd, A. C. *Rapp. Reun. Cons. int. Explor. Mer.* **172**, 137 (1978).
- Bruzhinin, A. D. & Pshenichny, B. P. in *Bioproductivity of Upwelling Ecosystems* (eds Barber, R. T. & Vinogradov, M. Y.) (Springer, Berlin, in the press).
- Dugdale, R. C. & Goering, J. J. *Anton Brunn Rep.* **4**, 53 (1970).
- Ryther, J. H., Menzel, D. W., Hulbert, E. M., Lorenzen, C. J. & Corwin, N. *Inv. Pesq.* **35**, 43 (1971).
- Walsh, J. J. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 497 (Plenum, New York, 1980).
- Whitledge, T. E. in *Upwelling Ecosystems* (eds Boje, R. & Tomczak, M.) 90 (Springer, Berlin, 1978).
- Dugdale, R. C. & Goering, J. J. *Limnol. Oceanogr.* **12**, 196 (1967).
- DeMaster, D. J. thesis, Yale Univ. (1979).
- Henrichs, S. M. thesis, MIT-WHOI (1980).
- Barber, R. T. & Smith, R. L. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 31 (Academic, New York, 1980).
- Cowles, T. J., Barber, R. T. & Guillen, O. *Science* **195**, 285 (1977).

# *In vivo* sequence requirements of the SV40 early promoter region

Christophe Benoist\* & Pierre Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

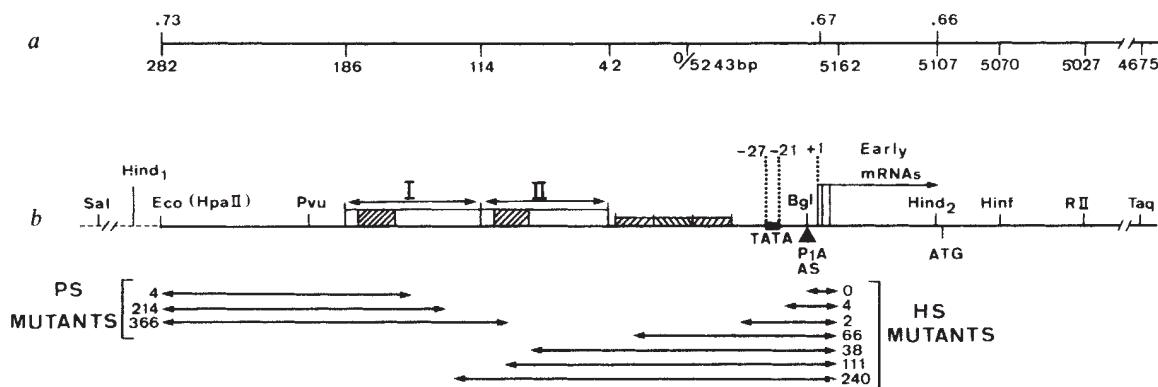
*To investigate the sequences necessary for proper initiation of transcription of SV40 early genes, we have constructed several deletion mutants in the promoter region. The TATA box region is apparently involved in fixing initiation precisely within a narrow area, but is dispensable for gene expression, while the sequences located more than 150 base pairs upstream are indispensable.*

In eukaryotic systems, the identity and even the existence of the DNA sequences which function as promoters for RNA polymerase B are as yet uncertain. DNA sequences of eukaryotic

genes (see refs 1, 2 for reviews) and their adjacent regions nevertheless contain an A+T-rich region, typically 5'TATA<sub>A</sub>A<sub>T</sub>3' at a strongly conserved distance 30 base pairs from the nucleotide coding for the first base of the mature mRNA, called the cap site and presumably the point at which transcription begins<sup>3</sup>. The A+T-rich region, known as the

\* Present address: Stanford University Medical School, Department of Genetics, Stanford, California 94305, USA.





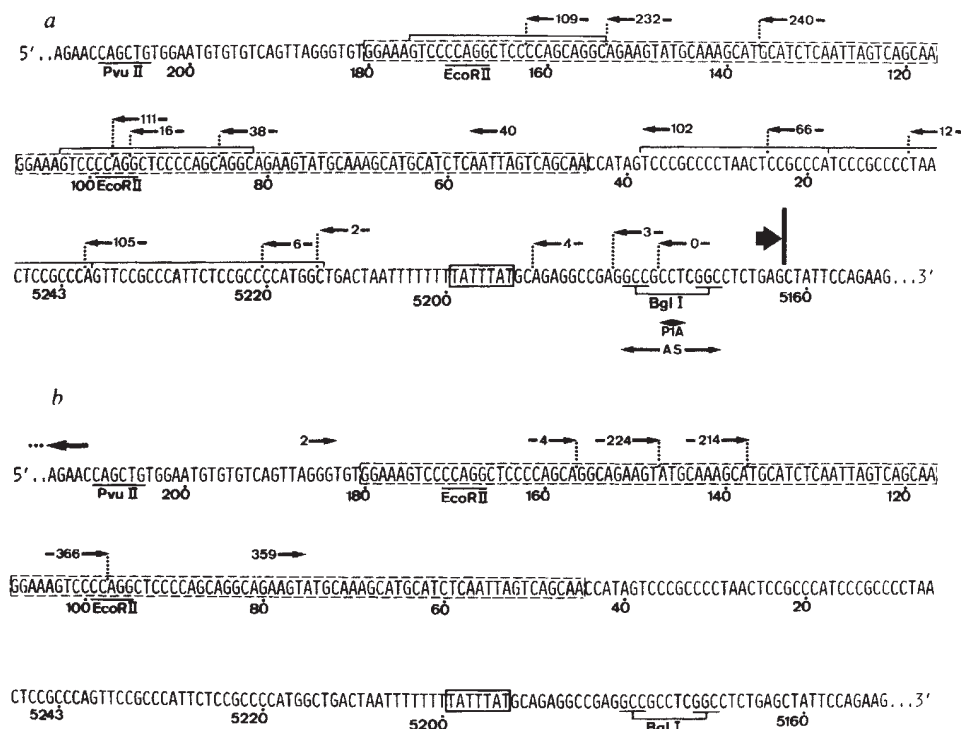
**Fig. 1** a, Scale in SV40 map units (above the line) or in base pairs (bp), the latter according to the numbering system<sup>13</sup> of Reddy *et al.*<sup>41</sup> modified to take into account a previously overlooked 17 base pairs<sup>29</sup>. b, Physical map of the SV40 early promoter region in the pSV1 recombinant<sup>13</sup>, and schematic representation of the deletions present in some of the mutants analysed in this study. SV40 sequences are represented by a solid line, pBR322 sequences by a dashed line. Sal, Hind, Eco, Pvu, Bgl, Hinf, RII and Taq represent sites for restriction enzymes SalI, HindIII, EcoRI, PvuII, BglI, HinfI, EcoRII and TaqI, respectively (only the relevant sites are shown). Eco (HpaII) indicates the EcoRI site of pSV1 created by the blunt-end ligation of the repaired SV40 HpaII site into the repaired EcoRI site of pBR322<sup>13</sup>. The position of the 5' ends of wild-type SV40 early mRNAs are shown by vertical lines, followed by a horizontal arrow which indicates the direction of transcription [+1 corresponds to the first cap site at position 5,172 (refs 19, 20)]. The position of ATG translation initiation codon is shown next to the restriction site Hind 2. The small solid box indicates the position of the TATA box (position -21 to -27 from position +1). Large boxes represent the 72-base pair repeats I and II, and the dashed areas within correspond to the G+C-rich blocks. The three small dashed boxes are the shorter, G+C-rich repeats (see Fig. 2a). The position of the deletions in mutants P1A and AS is shown under the BglI site. Horizontal arrows define the extent of the deletions, as determined by DNA sequence analysis, in representative HS and PS mutants. For the construction of P1A, pSV1 was linearized by partial cleavage with BglI (there are three BglI sites in the pBR322 portion of pSV1 along with the unique site of interest in the SV40 moiety). The BglI 3'-protruding extremities were removed with *Escherichia coli* DNA polymerase I (Biolabs), and the plasmids circularized by blunt-end ligation with T4 DNA ligase. P1A was isolated by screening recombinants for the absence of the BglI site in the SV40 sequences. The deletion in P1A led to the generation of a SacII site at this position. We took advantage of this unique site to cleave P1A by SacII, and, by the same procedure as that used for construction of P1A, isolated the mutant AS. To construct the mutants of the PS series, pSV1 was linearized at the unique EcoRI site, and treated with Bal 31 nuclease (BRL)<sup>13,42</sup> to remove up to 200 base pairs from the extremities thus generated. The DNA was then treated with SalI and ligated to an EcoRI-SalI fragment from pBR322. Selection of tetracycline-resistant bacteria allowed the isolation of mutants bearing deletions on the SV40 side of the EcoRI site only. The construction scheme of HS mutants has been reported in detail previously<sup>13</sup>.

Goldberg-Hogness or the TATA box, has been regarded as a promoter sequence analogous to the Pribnow box sequence of prokaryotic organisms. Nevertheless, genes (such as those coding for papovavirus late function and adenovirus-2 72,000 molecular weight (72K) DNA binding proteins)<sup>4-8</sup> lack the TATA box, so that its proposed function remains conjectural.

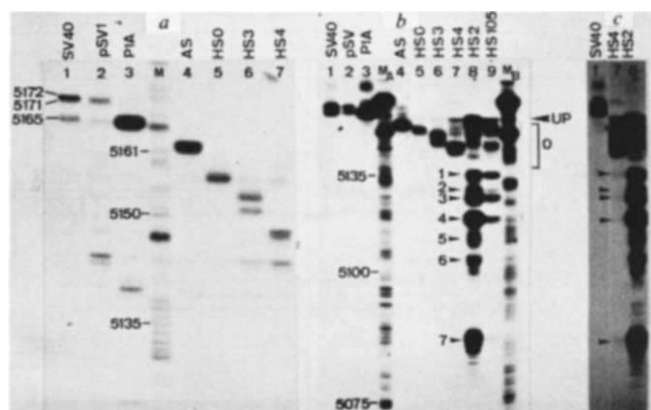
Using *in vitro* transcription systems that allow RNA polymerase B to transcribe cloned DNA templates<sup>9,10</sup>, the TATA

box has been shown to be involved in promoter function *in vitro*<sup>2,11,12</sup>, but *in vivo* studies suggest that these results should be extrapolated with caution. Thus for SV40 early genes, we and others<sup>13-15</sup> have shown that removal of all sequences upstream from the cap site, but not the mere removal of the TATA box<sup>13</sup>, abolishes expression. Similarly, the TATA box is not essential for expression of the sea urchin histone H2A gene injected into *Xenopus* oocytes although its absence leads to the use of several new cap sites<sup>16</sup>.

**Fig. 2** Sequence of the SV40 deletion mutants. The sequence corresponds to the DNA non-coding strand of the early promoter region which was cloned in pSV1, and is identical to that of wild-type SV40 strain 776<sup>29,41,43</sup>, except for an 8-base pair deletion (positions 116-108, see text). The numbering system is as in Fig. 1. The 72-base pair repeats are shown in two dashed boxes, the TATA box in a solid box. Horizontal lines above the sequence delimit the five G+C-rich blocks (see text). a, Sequence of mutants in the HS group, which all share a common end point at position 5,162 indicated by a vertical line and a large arrow. The other deletion end points are shown by thinner arrows (and dotted lines for those which have been sequenced). Deletions present in mutants P1A and AS around the BglI site are indicated by double pointed arrows. b, Sequence of the mutants in the PS series. All deletions originate at the EcoRI site (Fig. 1) and extend to the point defined by thin arrows. The extent of the deletions was established for all mutants by DNA sequencing except for mutants HS102, HS40, PS2 and PS359, where the size ( $\pm 8$  base pairs) of the deletions were estimated by gel electrophoresis after HindIII digestion. Sequences were obtained by the technique of Maxam and Gilbert<sup>21</sup> after 5'-end labelling at the HindIII site 2 (Fig. 1) and secondary cleavage with EcoRI (P1A, AS and HS mutants) or by end labelling at the HindIII site 1 and secondary cleavage with BglI (PS mutants). The exact sequence of some of the mutants is shown in Fig. 6a.



**Fig. 3** Mapping of the 5' ends of SV40 early mRNAs present in several rat (3T3) fibroblast transformed cell lines (see text). Cytoplasmic RNA was extracted from such cloned cell lines by NP40 lysis and phenol-chloroform extraction, and poly(A)<sup>+</sup> RNA selected by oligo(dT)-cellulose chromatography. **a**, Mapping by reverse transcriptase primer extension. A *HinfI*-*HindIII* primer (positions 5,070–5,107, Fig. 1) <sup>32</sup>P-5'-end labelled at its *HinfI* extremity, was purified by electrophoresis on an 8% acrylamide/7 M urea gel. It was then denatured in H<sub>2</sub>O for 2 min at 100 °C, quick-frozen in liquid nitrogen and thawed in 50 mM Tris-HCl pH 8, 10 mM MgCl<sub>2</sub> and 140 mM KCl in the presence of poly(A)<sup>+</sup> RNA (2–5 µg). After 5 min at 65 °C dATP, dGTP, dCTP and TTP (0.5 mM each) and β-mercaptoethanol (20 mM) were added and the samples were incubated for 2 h at 37 °C in the presence of 4 units of reverse transcriptase. Samples were diluted to 0.1 M NaOH, 3 mM EDTA and 5 M urea, and aliquots electrophoresed on 8% acrylamide/7 M urea thin gels<sup>44</sup>. RNAs were from cell lines transformed with wild-type SV40 or recombinant mutants as indicated. 'M' is a sequencing ladder obtained by Maxam and Gilbert<sup>21</sup> cleavage (A + C) of a *HinfI*-*EcoRI* fragment of pSV1 (positions 5,070–282) end labelled at its *HinfI* extremity. The positions of the 5'-end bands are given in Fig. 6 legend. Aside from the major cap sites, a few other bands can be seen (see lanes 2 and 3, for example) that probably correspond to reverse transcriptase 'false stops' as they are not obtained by S<sub>1</sub> mapping. **b**, S<sub>1</sub> nuclease mapping of the 5' ends of the RNAs (as indicated). The coding strand of SV40 fragment *EcoRII* G (positions 5,027–99) was 5'-end labelled and purified by polyacrylamide gel electrophoresis as described in ref. 21. After denaturation for 5 min at 85 °C, the probe was hybridized to RNAs (1–3 µg) from transformed cell lines for 12 h at 42 °C in 50% formamide, 10 mM PIPES pH 6.4 and 0.4 M NaCl. Digestion with S<sub>1</sub> nuclease (Miles) and subsequent steps were as in ref. 45. S<sub>1</sub>-resistant hybrids were analysed by electrophoresis on an 8% acrylamide/7 M urea thin gel. M<sub>A</sub> and M<sub>B</sub>: sequencing ladder from an *EcoRII*/*BglI* (5,027–5,177) cleaved by the technique of Maxam and Gilbert [M<sub>A</sub>, A + C; M<sub>B</sub>, C + T]. Arrows 1–7 indicate the common bands found with RNAs from all lines transformed with HS mutants lacking the TATA box (positions 5,134, 5,127, 5,124, 5,117, 5,112, 5,103 and 5,085). The 'O' bracket shows the area where initiation varies from one mutant to the next (see text). The 'UP' band corresponds to RNA species initiating upstream from the position of the deletion (see text). **c**, Same as **b**, where a longer exposure reveals the presence of bands 1–4 and 7 obtained with HS4 RNA (arrowheads), but not with SV40 RNA. Numbering of lanes as in **b**.



The SV40 early gene products (small t and large T antigens), which initiate viral DNA replication in permissive monkey cells, stimulate various host cell functions and transform non-permissive systems (see ref. 17 for review), are coded for by different mRNAs derived from a common primary transcript by different splicing patterns, and have identical 5' and 3' ends<sup>18,19</sup>. The major 5' ends of SV40 early mRNAs map at a few closely positioned cap sites, around positions 5,165 and 5,171 on the SV40 sequence (refs 19, 20; see also Figs 1, 6). Upstream from the cap sites are located, successively, a 27-base pair palindrome, a TATA box (TATTTAT, at position -21 to -27 base pairs from the most upstream of the cap sites), a nearly perfect triple tandem of a highly G + C-rich 21-base pair sequence, and a perfect tandem repeat of a 72-base pair sequence (-116 to -188 and -189 to -261 base pairs from the cap site) (see Figs 1, 2).

We have previously reported<sup>13</sup> that T-antigen synthesis can be induced by mutant plasmids containing an SV40 early region from which the TATA box has been deleted. We have now analysed the 5' ends of RNA species synthesized on such mutant templates, and conclude that initiation at the wild-type cap sites requires the presence of the TATA box. In the absence of the TATA box, transcription begins at several new sites. Moreover, we find that the expression of the SV40 early genes is severely impaired by deletions in the region of the 72-base pair repeats, which accordingly seems to have a crucial and independent role in initiation of transcription.

## Deletion mutants of SV40

The mutants described here were derived from the plasmid pSV1<sup>13</sup>, which contains the entire early gene segment (as the large *Bam*HI-*Hpa*II fragment of SV40), cloned into the *Eco*RI site of pBR322. From this, we have constructed three groups of deletion mutants (see Fig. 1 legend). After preliminary screening by restriction enzyme digestion, the extent of the deletions was established (in most cases, see Fig. 2) by DNA sequencing<sup>21</sup>.

The first group of mutants is represented by plasmids P1A and AS, with deletions between the TATA box and the cap sites of 3 and 9 base pairs, respectively (Figs 1, 2a). Figure 1 depicts schematically the position of the deletions in the second (PS) and third (HS) groups of mutants. Deletions in the PS series share an end point at the *Eco*RI site at the boundary of the SV40 and pBR322 sequences in pSV1, and extend progressively further downstream towards the early genes. The same pBR322 sequences are always present at the boundary, which greatly facilitates screening and eliminates one source of variability in

studying the effect of deletions in the SV40 sequence. Deletions in the HS mutants all extend upstream from position 5,162 (ref. 13), and thus delete the cap site and usually the TATA box, which is nevertheless left intact in HS0, HS3 and HS4 (Figs 2, 6). On mutant sequences, all have an 8-base pair deletion at the junction between the two 72-base pair repeated sequences (positions 116–108), apparently because the SV40 fragment cloned in pSV1 is derived from a variant of SV40 strain 776 (Fig. 2).

## In vivo expression of SV40 early genes in mutant plasmids

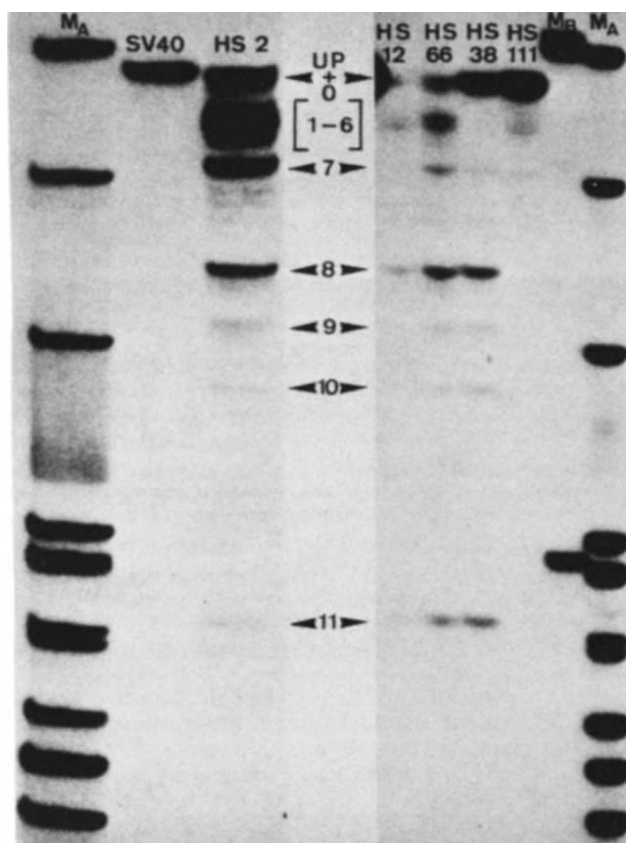
Our demonstration that deletion of all sequences upstream from the cap site abolishes SV40 early gene expression<sup>13</sup>, shows that some sequence element between the TATA box and *Hpa*II site

**Table 1** In vivo expression of the deletion mutants

| Short-term |                                                        | Long-term |                |
|------------|--------------------------------------------------------|-----------|----------------|
| Plasmids   | T antigen-positive cells per 2 × 10 <sup>4</sup> cells | Plasmids  | Foci per plate |
| pSV1       | 105                                                    | pSV1      | 132            |
| AS         | 72                                                     | HS4       | 115            |
| HS4        | 91                                                     | HS2       | 118            |
| HS2        | 78                                                     | HS105     | 72             |
| HS12       | 17                                                     | HS102*    | 17             |
| HS66       | 15                                                     | HS38*     | 21             |
| HS40       | 6                                                      | HS111*    | 7              |
| HS38       | 8                                                      | HS240*    | 1              |
| HS111      | 0                                                      | PS214*    | 15             |
| PS224      | 12                                                     | PS366*    | 13             |
| PS366      | 0                                                      |           |                |

Short-term experiments: Plasmid DNA (0.2 µg) was diluted in 200 µl of Dulbecco's modified Eagle's medium (DMEM), followed by the addition of 200 µl of a 1 mg ml<sup>-1</sup> solution of DEAE-dextran (Sigma) in DMEM. Sub-confluent CV1 monkey cells, grown in 60-mm Petri dishes<sup>13</sup>, were washed twice with phosphate-buffered saline (PBS); the DNA solution was added and, after incubation for 15 min at room temperature, the cells were washed once with PBS and once with complete medium. After addition of medium, they were incubated at 37 °C. Indirect immunofluorescence was performed 40 h after transfection, essentially as described in ref. 39, 10 µl of the antibodies<sup>13</sup> being spotted directly on to the cell monolayer fixed to the dish with acetone/methanol (3:7 v/v). T antigen-positive cells were counted under a fluorescence microscope. Long-term experiments: Transformation of non-permissive cells. Fisher rat fibroblasts (FR3T3), grown in 100-mm Petri dishes<sup>13</sup>, were transfected with 1 µg of plasmid DNA, as described in ref. 40. Dense foci overgrowing the cell monolayer were counted 3 weeks after transfection. Asterisks indicate that foci were counted 4 weeks after transfection and that they were checked by immunofluorescence to eliminate foci due to spontaneous transformation (5–10 per dish). All the data shown in this table represent results obtained in typical experiments. The absolute figures are only indicative, however, as they may vary by as much as 50% between experiments.





**Fig. 4**  $S_1$  nuclease mapping showing the full extent of the 5' ends of RNAs from cells transformed with HS mutants lacking the TATA box. The experimental protocol is the same as that described in Fig. 3b legend (4  $\mu$ g of poly(A)<sup>+</sup> RNA for SV40 and HS2; HS12, 8  $\mu$ g; HS66, 38  $\mu$ g; HS111, 15  $\mu$ g); the probe used was the coding strand of the *TaqI*-*BglI* fragment of SV40 (positions 4,675–5,180), 5'-labelled at its *TaqI* extremity.  $S_1$ -resistant hybrids were analysed on a 5% acrylamide/7 M urea thin gel.  $M_A$ : 5'-end labelled fragments from a *HpaII* digest of pBR322 (sizes from top in base pairs: 527, 404, 309, 242, 238, 217, 201, 190, 180).  $M_B$ : 5'-end labelled fragments from a *HinfI* digest of SV40. The origin of the RNAs is indicated at the top of the lanes. The resolution on a 5% polyacrylamide gel is not sufficient to separate the band mapping at the position of the deletion (UP) from those immediately below (region O) (see Fig. 3b). Similarly, bands 1–6 are poorly separated. Bands 7–11 correspond to 5' ends of RNA species mapping at positions ( $\pm 10$  base pairs) 5,085, 5,023, 4,985, 4,962 and 4,897, respectively. (Only the major bands 1–11 are numbered, but a large number of more minor cap sites are also present, for example, HS2 RNA, where at least 10 such bands can be seen between bands 7 and 11 on the original autoradiogram.)

(position 282, Fig. 1) is essential for transcription. We have therefore analysed the *in vivo* expression of the early genes of HS and PS mutants (Fig. 2).

Two approaches were followed (Table 1). In short-term experiments, gene expression was detected by T antigen-specific immunofluorescence 40 h after transfection of permissive CV1 monkey cells with the recombinant plasmids. In long-term experiments, we assayed the ability of the deletion mutants to induce morphological transformation of rat 3T3 fibroblast cells (tested by the appearance of T antigen-positive dense foci overgrowing the cell monolayer).

In the short-term experiments, HS mutants (Fig. 2a) up to and including HS38 induced a T antigen-positive immunofluorescence pattern, with an intensity unaffected by deletion of the TATA box (HS2 and 6), but inversely proportional to the length of the deletions (HS66, 40 and 38). On the other hand, the responses with HS16, 111, 240, 232 and 109 were not unambiguously positive. With the sequence of PS mutants (Fig. 2b), gene expression was allowed down to those in PS224 and 214, but with these two mutants the fluorescence intensity was lower than that obtained with the wild-type recombinant pSV1; in contrast, PS366 and 359 gave negative results. We have also assayed by blot hybridization<sup>22</sup> for SV40-specific early RNAs in

poly(A)<sup>+</sup> RNA extracted from CV1 cells 40 h after transfection with various plasmids, and have found none in cells when the mutants were negative by immunofluorescence (not shown). The sensitivity of such experiments is low, however, and it is possible that a low level of transcription (<10% of that obtained with pSV1) occurs with mutants which do not induce T-antigen immunofluorescence.

These experiments therefore show that the region of the 72-base pair repeats (Figs 1, 2) is necessary for efficient expression of the early genes. However, morphological transformation studies were somewhat more ambiguous. As expected, rat 3T3 cells were transformed by HS mutants up to HS38, but the efficiency of transcription was decreased with mutants such as HS40 or HS38. Using RNA blot analysis, we have quantitated the amount of early RNAs present in lines of such transformants. Like Gluzman *et al.*<sup>14</sup>, we found that the levels of these RNAs were higher in cells transformed with mutants bearing deletions downstream from the TATA box than in cells transformed with pSV1 (although the increase varied from 2- to 10-fold with different lines; our unpublished results). These levels were not affected by further deletion of the TATA box, but decreased with HS mutants bearing longer deletions (not shown). However, in contrast to the short-term experiments, T antigen-positive transformed foci were induced by HS111, HS240 and PS366, although in reduced numbers (5–10 foci per  $\mu$ g of DNA per  $10^6$  cells compared with 50–100 for pSV1 and short deletion HS mutants) and more slowly (25–30 days compared with 10–14 days).

### Transcription initiation on mutants retaining the TATA box

Using both  $S_1$  nuclease mapping with single-stranded <sup>32</sup>P-5'-end-labelled probes<sup>23</sup> and reverse transcriptase primer extension mapping<sup>19</sup> (see Figs 1, 3), we have investigated the 5' ends of SV40-specific RNAs present in cloned rat fibroblast cell lines transformed by HS mutants. We believe both types of experiments to be necessary because of the risk of artefacts, particularly with the primer extension method (see Fig. 3 legend). We have assigned (Fig. 6a) the major 5' ends of SV40 early mRNAs to positions 5,172, 5,171 and 5,165 (both in RNA isolated 32 h after infection by wild-type virus<sup>24</sup> and in lines transformed by wild-type SV40 or pSV1; Fig. 3a and unpublished results). However, because results obtained by the two methods often differ by 1 or 2 base pairs (compare, for instance, Fig. 3a with Fig. 5a of the accompanying paper<sup>24</sup>), and because of discrepancies in previous reports (refs 18–20 and Y. Groner, personal communication), we do not consider these assignments definitive.

Figure 3a and b give the 5'-end analysis of RNAs transcribed from deletion mutants retaining the TATA box. The 3-base pair deletion in P1A results in initiation mainly at position 5,165, with almost complete disappearance of the two wild-type 5' ends upstream at positions 5,172 and 5,171. On mutants with longer deletions downstream from the TATA box (AS, HS0, HS3, HS4), transcription begins further down, with the 5' ends always mapping at about the same distance from the TATA box (27–34 base pairs from the first T, Fig. 6a). This shift seems to be independent of the presence (P1A, AS) or absence (HS0, HS3, HS4) of the wild-type cap sites. In those mutants which do conserve the TATA box, there are very few cap sites.

### Transcription initiation on mutants lacking the TATA box

The clearest consequence of the removal of the TATA box is the generation of multiple cap sites, as shown for mutants HS2 and HS105 (Fig. 3b). Identical bands (arrows 1–7 in Fig. 3b) were observed with all the mutants lacking the TATA box, but with varying intensities (see below). Between band 1 and that corresponding to the end point common to all deletions (arrowhead 'UP'), patterns varied between mutants (region O in Fig. 3b) although some cap sites were common to several mutants (Fig.

5). We investigated this multiplicity further by  $S_1$  mapping using a single-stranded probe end labelled at the *TaqI* site (position 4,675, Fig. 1), and found that in the absence of the TATA box, initiation occurs up to 270 base pairs downstream from the normal 5' ends (Fig. 4, band 11). These cap sites are not usually seen with wild-type SV40 RNA (Fig. 4), but it is just possible to detect them using very large amounts of such RNA (not shown). As the probes used in the experiments shown in Figs 3b and 4 were derived from wild-type SV40 DNA, RNA species from HS mutants and with their 5' ends mapping upstream from the position of the deletion lead to the appearance of a band ('UP' in Figs 3b, 4) at that position. With mutants such as HS2 and 105 these RNAs are relative minor, with most of the initiation occurring downstream.

Figure 4 shows the effect of increasingly long HS deletions including the TATA box and sequences upstream. Although identical bands are found with mutants HS2, 12, 66 and 38, the ratio of the intensity of these bands varies, the lower bands becoming more prominent with increasing deletion. For example, whereas bands 1–6 are most important in HS2 RNA, they are barely detectable in HS38 RNA, while the relative intensities of the lower bands 8–11 increase (see Fig. 6b for quantitation). It thus appears (similar results were obtained reproducibly in several experiments) that, although identical cap sites are shared by various HS mutants, their use is influenced by their distance from sequences found upstream from the HS38 deletion end point.

HS111 RNA shows a very different pattern, with the disappearance of the lower bands 8–11 and a relative reappearance of the higher bands 1–7; most of the RNA initiates upstream from the position of the deletion ('UP' in Fig. 4). This 'UP' band is also intense in HS38 RNA. We investigated the cap sites located upstream from the position of the deletion by mapping the 5'

ends of RNAs from cell lines transformed with HS2, 105, 66 (not shown), 38 and 111 (Fig. 5) with probes derived from corresponding mutant DNAs. Autoradiograms from experiments such as those shown in Figs 3, 4 and 5, were scanned and the results are summarized in Fig. 6b. With HS2, 105 and 66 RNAs the 'upstream' RNA species are heterogeneous and quite minor, their 5' ends mapping immediately upstream from the position of the deletion. For HS111, however, they are of much higher relative abundance, and are represented by a set of bands (Fig. 5) which correspond to initiation within the 72-base pair repeat I (Figs 1, 2). This set of bands is also observed, but at much lower relative abundance (Fig. 6b) and with a somewhat different distribution of intensities (Fig. 5), with HS38 RNA, where most of the upstream RNAs map far into the late region (X on Figs 5, 6b). Although very faint bands, perhaps corresponding to the set observed with HS111 RNA, were detected in HS2 RNA, no other HS RNAs showed any species equivalent to the 'X' bands seen with HS38 RNA.

Although we have assumed throughout that the bands obtained in  $S_1$  mapping experiments correspond to 5' ends of the RNAs it is possible that these bands correspond to the position of acceptor sites in RNA splicing. However, multiple splicing events would then appear suddenly after the small deletion from HS4 to HS2, and the relative use of acceptor sites would be further modified by increasing deletions. Splicing can be ruled out for RNAs transcribed on mutants which conserve the TATA box and for bands 1–4 in the cases where this sequence is absent, as these 5' ends are also found by reverse transcriptase primer extension mapping (Fig. 3a and data not shown). Another interpretation is that as the RNAs studied are extracted from transformed cell lines, some of the bands seen in the  $S_1$  nuclease mapping experiments could have been due to sequence rearrangements or integration behind cellular promoters during transformation. As such events are unlikely to occur similarly in a number of cell lines transformed by different mutants, this objection is probably not valid for bands 1–11 which are found with all mutants, nor for the set of bands found in both HS111 and HS38 RNAs, but may explain the bands found only with HS38 RNA, in region X (Fig. 5).

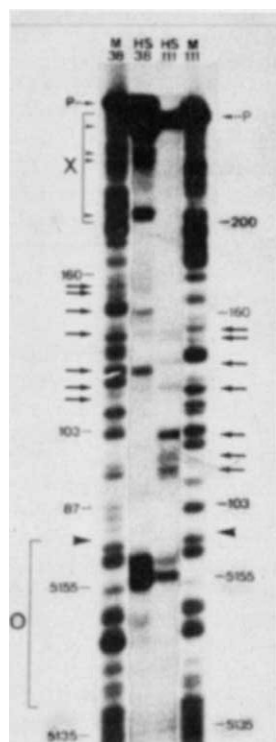
## Function of the TATA box region

Our studies on the above mutants, in agreement with recent reports (refs 14, 16 and Y. Gluzman, personal communication), clearly show that the TATA box has a key role in fixing the start sites within a narrow area (we assume that the RNA 5' ends correspond to transcription start sites). Deletions in such mutants shift these sites so that initiation always occurs 27–34 base pairs downstream from the first T of the TATTAT sequence (the slight variations in this distance and in the number of start sites presumably reflect the influence of the actual DNA sequence around the cap site<sup>2,24</sup>).

When the TATA box is absent, new cap sites appear, but no TATA-like sequences are found upstream from the cap sites (except possibly for band 10, Fig. 4). We hypothesize that sequence elements, not obviously related to the canonical TATA box sequence, can functionally substitute for the TATA box to dictate the position of RNA 5' ends and might also direct initiation 30 base pairs downstream, as the first 5' end common to all TATA-minus mutants (band 1 in Figs 3b, 6) maps 30 base pairs downstream from the position of the deletion. In the wild-type situation, these substitute elements would be masked by the TATA box, which might be recognized more readily by RNA polymerase and/or other factors involved in transcription initiation. A functionally similar element might be operative in polyoma late genes and the gene for the adenovirus 72K DNA binding protein, which exhibit a microheterogeneity of mRNA 5' ends over a narrow area, but do not have TATA-like sequences at the expected position<sup>4,6</sup>.

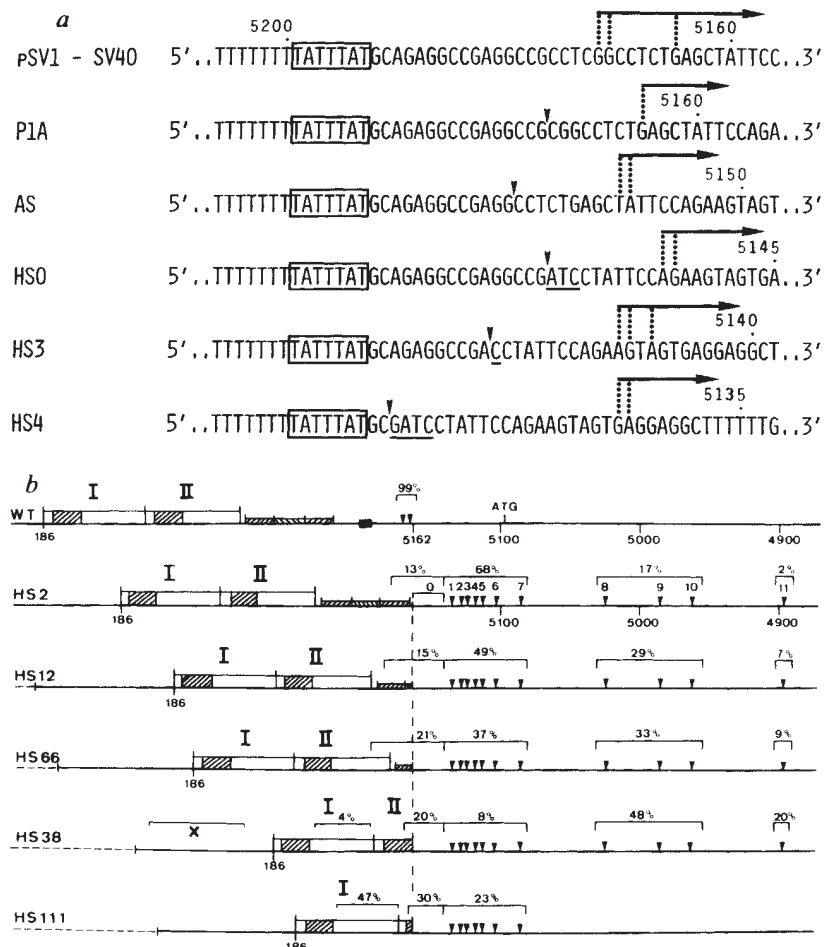
The key sequences performing the 'fixing' function for initiation at the major early cap sites must be present in HS4, but absent from HS2 (Figs 2, 3). Although the TATA box is included in this region, other sequences could influence its

**Fig. 5**  $S_1$  nuclease mapping of 5' ends of RNA species that initiate upstream from the position of the deletion, in RNA extracted from cell lines transformed with HS38 and HS111. The probes used were the *HinfI*-*EcoRI* fragments (5,070–282) derived from HS38 and HS111 DNAs, and 5'-end labelled at their *HinfI* extremity. The specific activities of the probes were different, as were the amounts of RNA used (HS38, 2  $\mu$ g; HS111, 3.5  $\mu$ g).  $M_{38}$  and  $M_{111}$  are sequence ladders obtained by (A + C) cleavage of the probes. HS38 and HS111: the probes were hybridized to RNA from corresponding cell lines, digested with  $S_1$  nuclease and electrophoresed as described in Fig. 3 legend. The markers and experimental tracks are from different exposures of the same gel. Positions are those on the wild-type SV40 sequence (numbering as in Figs 1 and 2) and arrowheads indicate the position of the deletion (5,162). 'O' brackets the region 0 (defined in Fig. 3 legend) between the position of the deletion and band 1 at position 5,135. Arrows on the right indicate the bands obtained with HS111 RNA (positions 117, 119, 120, 135, 144, 151 and 153). These bands are observed at identical positions with HS38 RNA (arrows on the left of the figure; the bands are faint but were clearly visible on the original autoradiogram). Bands obtained with HS38 RNA only are bracketed and labelled X. In both HS38 and HS111, a band can be seen which corresponds to full-length unreacted probe (dotted arrow 'P'). That this band is not due (at least for HS111) to RNA species initiating further upstream was established by an experiment in which the *HinfI* end-labelled probe extended for 600 base pairs into the pBR322 sequences (Fig. 1). No additional bands were observed in this experiment (not shown).





**Fig. 6** Location of the 5' ends of early RNAs present in lines transformed by deletion mutants of pSV1. *a*, Position of cap sites on mutants that retain the TATA box. The sequences of the mutants are shown, aligned on the 'TATTTAT' sequence (boxed). Arrowheads indicate the position of the deletion; the few extra nucleotides which originate from the construction of HS mutants are underlined<sup>13</sup>. Symbols for cap sites are as in Fig. 1. Their position was derived from the results shown in Fig. 3a and other unpublished S<sub>1</sub> and reverse transcriptase mapping experiments and map at the following position: P1A, 5,165; AS, 5,160 and 5,161; HS0, 5,154 and 5,155; HS3, 5,149, 5,151 and 5,152; HS4, 5,145 and 5,146. As discussed in the text these positions are only tentative ( $\pm 2$  base pairs). Only the major cap sites are represented, yet, as can be seen in Fig. 3a, minor cap sites exist; for example, the HS0 5' ends are also present in AS, HS3 and HS4 RNAs. *b*, Position and relative contribution of the RNA 5' ends found with HS mutants lacking the TATA box. Symbols are the same as in Fig. 1, as is the numbering system. The common deletion end point is shown by a vertical dashed line. Cap sites are shown by arrowheads and are numbered (on the second line only) as defined in Figs 3b and 4. The cap sites that vary from one mutant to the next ('region O', and the region immediately upstream from the deletion end point) are not shown individually, nor the closely spaced cap sites of mutants HS38 and HS111 that map within the 72-base pair repeat I. The brackets and the percentages above the cap sites indicate the proportion of RNA species initiating in that region, as deduced from scans of autoradiograms from several S<sub>1</sub> mapping experiments with wild-type probes end labelled at the *Hin*II, *Eco*RII, *Taq*I sites (Fig. 1) or experiments with mutant probes as shown in Fig. 5. Percentages for HS38 are calculated omitting the bands in region X, which represent 45% of HS38 RNA 5' ends.



function. Indeed, bands 1–4 and 7, characteristic of mutants lacking the TATA box, can be detected faintly in HS4 RNA (Fig. 3c), but are absent in RNA transcribed on the wild-type gene or on HS mutants with shorter deletions. It is as if the TATA box in HS4 were beginning to relinquish its dominance, an interpretation supported by the relative inefficiency of HS4 DNA (see accompanying article<sup>24</sup>) in directing accurate initiation of transcription *in vitro* to a site about 30 base pairs downstream.

The deletion of the TATA box does not seem to affect early gene expression quantitatively (Table 1). The possibility that down mutations would be masked if they eliminated T-antigen binding and repression<sup>25–28</sup> probably does not apply as the T-antigen binding site II is already absent from HS4<sup>26,28</sup>.

### Far upstream sequences are required for transcription initiation

The most important element for SV40 early gene expression is clearly not the TATA box, for mutant PS366, which has a deletion far upstream from the TATA box, transforms rat cells very inefficiently and does not induce detectable T antigen in short-term experiments, while the expression of various mutants in monkey cells indicates that, at least in this system, the essential elements lie in the region of the 72-base pair repeats. Results obtained with mutants HS38 and PS214 suggest that deletion of one of these two repeats still allows transcription, in keeping with studies of mutants or variants of SV40 which possess only one copy of this repeat (see ref. 29 for refs). Although the lower efficiency of mutants HS66 or HS38 (Table 1) might suggest the involvement of other sequences, such as the G+C-rich 21-base pair repeats, this could also be accounted for by the position of the 5' ends of RNAs synthesized on such templates, downstream from the normal T-antigen initiation codon. The results obtained with the PS214 and PS224 are in

contrast with the aforementioned mutants and variants of SV40, whose early gene expression appears normal. This might indicate the involvement of sequences even further upstream than the 72-base pair repeats, or that pBR322 sequences have some inhibitory effect when brought closer to the gene.

Increasingly long deletions (HS2, HS12, HS66, HS38) affect the relative intensities of the common bands 1–11 with a pattern (Fig. 6b), suggesting that initiation takes place predominantly at a fixed distance (broadly speaking 150–200 base pairs) from sequences located upstream from the deletion endpoint of HS38. This pattern is drastically altered with mutant HS111, which bears a deletion only 12 base pairs longer than that of HS38. In addition, HS111 does not induce detectable T-antigen synthesis in CV1 cells, whereas HS38 does. There are five sets of highly G+C-rich 20-base pair sequences in the early SV40 promoter region (Figs 1, 2). The G+C-rich sequence of the 72-base pair repeat II is almost eliminated in HS111, but essentially left intact in HS38. We suggest that this sequence is involved in controlling the pattern of initiation in mutants HS2 to HS38; HS111 transcription would then be under the control of the same G+C-rich sequence of repeat I. If this is the case, the effect of the G+C-rich sequences must be complex: predominant initiation occurs in HS111 only 10–60 base pairs downstream from the G+C-rich sequence of repeat I, whereas it occurs ~200 base pairs downstream from its counterpart in repeat II, when this is present (HS2 to HS38). The relatively constant amount of RNA species initiating in region O or immediately upstream (Fig. 6b) is also puzzling.

We have observed that mutants which do not support early gene expression in CV1 cells, such as HS111, HS240 and PS366, can, albeit inefficiently, nevertheless transform rat fibroblast cells. As these transformants could reflect a low level of transcription, not detected in short-term experiments because of a lower sensitivity, we are now investigating the extent of sequences necessary for such transformation (bearing in mind

the possibility that template rearrangements could be involved; see above).

We do not know how sequences lying more than 150 base pairs from the cap sites exert their influence. The structure of SV40 chromatin is known to be in a very 'open' conformation in the early promoter region<sup>30-34</sup>. Moreover, the chromatin structure around the 5' end of several 'active' cellular genes is extremely sensitive to DNase I (refs 35, 36 and M. Bellard, personal communication). It is possible that the upstream sequences somehow induce this peculiar chromatin structure, which in turn would allow the RNA polymerase to interact, directly or indirectly, with the TATA box. Alternatively (or concomitantly), these sequences could be part of the RNA polymerase binding site<sup>24</sup>, or interact with proteins exerting a positive control over transcription.

Can the organization found for the SV40 early promoter region be generalized to other systems? There is no clear homology between the SV40 sequence upstream from the TATA box and regions upstream from the 5' end of other genes. However, both polyoma early genes (R. Kamen, personal communication) and the sea urchin H2A gene (R. Grosschedl and M. Birnstiel, personal communication) have sequences lying upstream from the TATA box, in a position roughly analogous to that of the 72-base pair repeats in SV40, which are essential

for gene expression. Thus, it is possible that sequence elements will be found in other genes that have an analogous role to that of the 72-base pair repeat region in SV40.

In conclusion, if a eukaryotic promoter region is defined as being composed of all the sequences necessary for proper initiation of transcription *in vivo*, the SV40 early promoter has (at least) three components: the TATA box and cap site regions which accurately position initiation of transcription, and the sequences in the region of the 72-base pair repeats which allow, within the area they potentiate, initiation of transcription and TATA box function. The requirement for the far upstream sequences, which have no counterpart in prokaryotes, may reflect the importance of chromatin structure in the control of gene expression in eukaryotes<sup>37,38</sup>.

We thank Drs J. Beard and R. Wilsnack for gifts of reverse transcriptase and antiserum, Drs R. Breathnach, K. O'Hare and D. Mathis for helpful advice, K. Dott, J. M. Garnier, M. C. Gesnel, L. Thia-Toong and Dominique Schmitt for technical assistance, and B. Boulay for photography. This work was supported by grants from the INSERM (ATP 72.79.104, CL 801029 and CL 801033), the CNRS (ATP 'Biologie Moléculaire du Gène'), the Fondation Française pour la Recherche Médicale and from the Fondation S. and C. Del Duca.

Received 1 October 1980; accepted 22 January 1981.

- Breathnach, R. & Chambon, P. A. *Rev. Biochem.* (in the press).
- Corden, J. *et al. Science* **209**, 1406-1414 (1980).
- Ziff, E. & Evans, R. *Cell* **15**, 1463-1475 (1978).
- Flavell, A., Cowie, A., Arrand, J. & Kamen, R. *J. Virol.* **33**, 902-908 (1980).
- Haegeman, G. & Fiers, W. *Nucleic Acids Res.* **5**, 2359-2371 (1978).
- Baker, C., Herisse, J., Courtois, G., Galibert, F. & Ziff, E. *Cell* **18**, 569-580 (1979).
- Groner, Y., Carmi, P. & Aloni, Y. *Nucleic Acids Res.* **4**, 3959-3968 (1978).
- Ghosh, P., Reddy, V., Swinscoe, J., Lebowitz, P. & Weissman, S. *J. molec. Biol.* **126**, 813-846 (1979).
- Weil, P., Luse, D., Segall, J. & Roeder, R. *Cell* **18**, 469-484 (1979).
- Manley, J., Fire, A., Cano, A., Sharp, P. & Gelfand, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855-3859 (1980).
- Wasylyk, B., Keding, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367-373 (1980).
- Wasylyk, B. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Benoist, C. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3865-3869 (1980).
- Gluzman, Y., Sambrook, J. & Frisque, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3898-3902 (1980).
- Mueller, C., Graessmann, A. & Graessmann, M. *Cell* **15**, 579-586 (1978).
- Grosschedl, R. & Birnstiel, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432-1436 (1980).
- Acheson, N. in *Molecular Biology of Tumor Viruses Pt 2* (ed. Tooze, J.) 125-204 (Cold Spring Harbor, New York, 1980).
- Thompson, J., Radonovich, M. & Salzman, N. *J. Virol.* **31**, 437-446 (1979).
- Reddy, V., Ghosh, P., Lebowitz, P., Piatak, M. & Weissman, S. *J. Virol.* **30**, 279-296 (1979).
- Haegeman, G. & Fiers, W. *J. Virol.* **35**, 955-961 (1980).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-580 (1980).
- Alwine, J., Kemp, D. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Weaver, R. & Weissmann, C. *Nucleic Acids Res.* **7**, 1175-1193 (1979).
- Mathis, D. J. & Chambon, P. *Nature* **290**, 310-315 (1981).
- Reed, S., Stark, G. & Alwine, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3083-3087 (1976).
- Tjian, R. *Cell* **13**, 165-179 (1978).
- Shalloway, D., Kleinberger, T. & Livingston, D. *Cell* **20**, 411-422 (1980).
- Myers, R. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6491-6495 (1980).
- Van Heuverswyn, H. & Fiers, W. *Eur. J. Biochem.* **100**, 51-60 (1979).
- Scott, W. & Wigmore, D. *Cell* **15**, 1511-1518 (1978).
- Sundin, O. & Varshavsky, A. *J. molec. Biol.* **132**, 535-546 (1979).
- Varshavsky, A., Sundin, O. & Bohn, M. *Cell* **16**, 453-466 (1979).
- Saragosti, M., Moyné, G. & Yaniv, M. *Cell* **20**, 65-73 (1980).
- Wigmore, D., Eaton, R. & Scott, W. *Virology* **104**, 462-473 (1980).
- Stalder, J. *et al. Cell* **20**, 451-460 (1980).
- Wu, C. *Nature* **286**, 854-860 (1980).
- Flavell, R. *Nature* **285**, 356-357 (1980).
- Mathis, D., Oudet, P. & Chambon, P. *Prog. Nucleic Acids Res. molec. Biol.* **24**, 1-55 (1980).
- Mertz, J. & Berg, P. *Virology* **62**, 112-124 (1974).
- Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725-731 (1978).
- Reddy, V. *et al. Science* **200**, 494-502 (1978).
- Legerski, R., Hodnett, J. & Gray, H. *Nucleic Acids Res.* **5**, 1445-1464 (1978).
- Fiers, W. *et al. Nature* **273**, 113-120 (1978).
- Sanger, F. & Coulson, A. *FEBS Lett.* **87**, 107-110 (1977).
- Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* **65**, 718-749 (1980).

# The SV40 early region TATA box is required for accurate *in vitro* initiation of transcription

Diane J. Mathis\* & Pierre Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

*The SV40 early region TATA box functions in vitro to fix initiation of transcription within a narrow region identical to that in vivo. In the absence of the TATA box region initiation occurs at multiple specific sites, as in vivo. However, certain sequences far upstream, crucial for efficient in vivo expression, are not essential in vitro.*

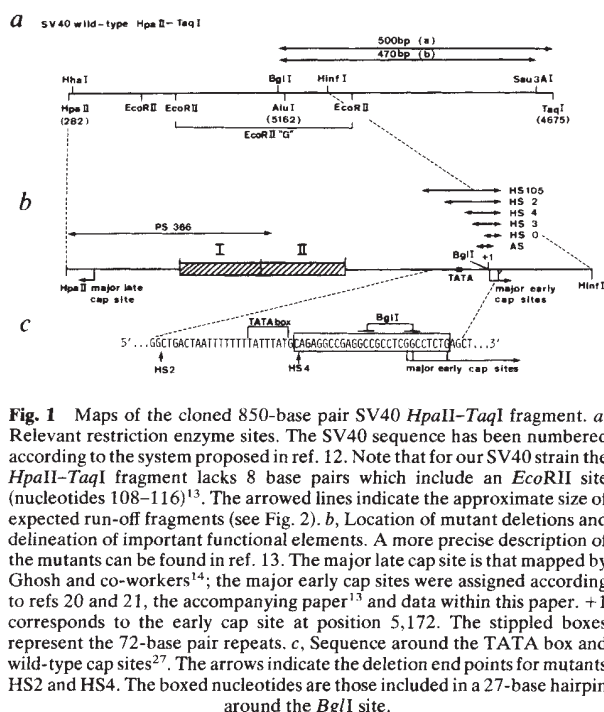
THE development of cell-free systems allowing selective and accurate initiation of transcription by RNA polymerase B (II)<sup>1,2</sup> has given a new impetus to studies of the control of eukaryotic gene expression. Viral<sup>1-3</sup> and cellular<sup>3,4</sup> genes have been transcribed using either a whole-cell<sup>2</sup> or cytoplasmic (S100)<sup>1</sup> extract

from HeLa or KB cells, and have generally shown close correspondence between the *in vitro* transcription start site for a given gene and its *in vivo* cap site (the nucleotide coding for the 5'-end of the mature mRNA).

This apparent fidelity has allowed the use of a HeLa cell-free system to delineate the DNA sequences required for specific RNA polymerase B initiation. Corden and co-workers<sup>5</sup> have reported that the DNA segment 12-32 nucleotides upstream

\* Present address: Stanford University School of Medicine, Department of Medical Microbiology, Stanford, California 95304, USA.





**Fig. 1** Maps of the cloned 850-base pair SV40 *HpaII*-*TaqI* fragment. *a*, Relevant restriction enzyme sites. The SV40 sequence has been numbered according to the system proposed in ref. 12. Note that for our SV40 strain the *HpaII*-*TaqI* fragment lacks 8 base pairs which include an *EcoRII* site (nucleotides 108–116)<sup>13</sup>. The arrowed lines indicate the approximate size of expected run-off fragments (see Fig. 2). *b*, Location of mutant deletions and delineation of important functional elements. A more precise description of the mutants can be found in ref. 13. The major late cap site is that mapped by Ghosh and co-workers<sup>14</sup>; the major early cap sites were assigned according to refs 20 and 21, the accompanying paper<sup>13</sup> and data within this paper. +1 corresponds to the early cap site at position 5,172. The stippled boxes represent the 72-base pair repeats. *c*, Sequence around the TATA box and wild-type cap sites<sup>27</sup>. The arrows indicate the deletion end points for mutants HS2 and HS4. The boxed nucleotides are those included in a 27-base hairpin around the *BglI* site.

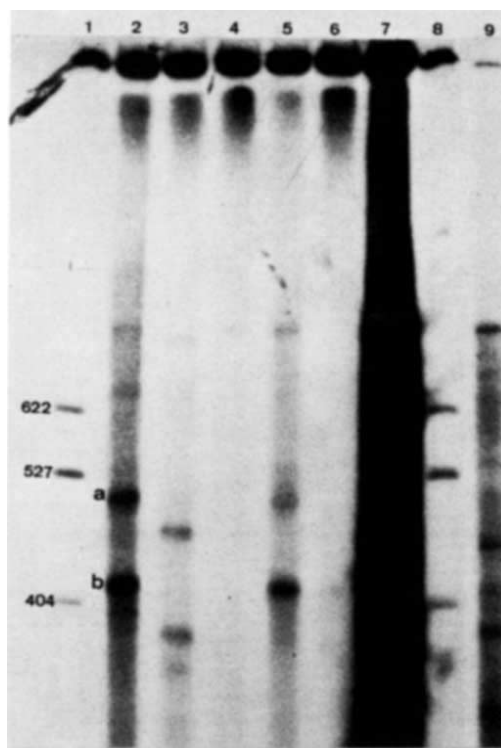
from the cap site of the adenovirus 2 (Ad-2) major late transcription unit suffices to promote specific initiation *in vitro*; similar results were obtained with a DNA segment 10–44 base pairs upstream from the conalbumin gene cap site<sup>5</sup>. These segments share the sequence 5'-CTATAAAAGGGG-3', which includes the TATA or 'Goldberg-Hogness' box (underlined) thought to be involved in promoting specific initiation, because it is conserved before the 5' end of most RNA polymerase B-transcribed genes<sup>6–9</sup>. Moreover, specific RNA polymerase B initiation is essentially abolished on mutants with deletions involving the TATA sequence<sup>5</sup>, and when the conalbumin gene TATAAAA sequence is mutated to TAGAAAA, specific *in vitro* transcription is drastically reduced<sup>10</sup>. The conclusion, then, is that the TATA box has a fundamental role in promoting specific initiation by RNA polymerase B, at least *in vitro* using naked DNA templates and HeLa cell-free extracts.

Two studies cast doubt on whether the TATA box has an identical role *in vivo*. Deletion of the TATA box does not abolish transcription of the sea urchin H2A histone gene in *Xenopus laevis* oocytes<sup>11</sup> nor that of the SV40 early region in mammalian cells<sup>12,13</sup>, but instead generates a new set of start sites. One might argue, of course, that these genes have atypical promoters—the histone genes are preceded by extensive A + T-rich 'spacer' regions and the SV40 early region TATA box is embedded in a highly A + T-rich region (5'-TAATTTTTTTT-TATTTAT-3'). Such sequence organizations are not typical of known RNA polymerase B-transcribed genes<sup>9</sup>, and the H2A genes and SV40 early region TATA boxes may not have retained their usual function(s). Alternatively, RNA polymerase B may not see the TATA box in the same context *in vivo* as *in vitro*.

It seems necessary, then, to assess the fidelity of the *in vitro* transcription systems at another level. Do mutations within a given DNA segment affect its *in vitro* and *in vivo* transcription identically? Here we compare the SV40 early region transcripts generated in a HeLa S100 extract with those detected *in vivo* in mammalian cells. We first establish that *in vitro* RNA polymerase B initiates transcription accurately on the early gene region, then compare the effects of various deletions in the promoter region on initiation *in vivo* and *in vitro*. The construction of these mutants and a detailed study of their *in vivo* expression are described in the accompanying article<sup>13</sup>.

## RNA polymerase B can initiate accurately on the SV40 early region *in vitro*

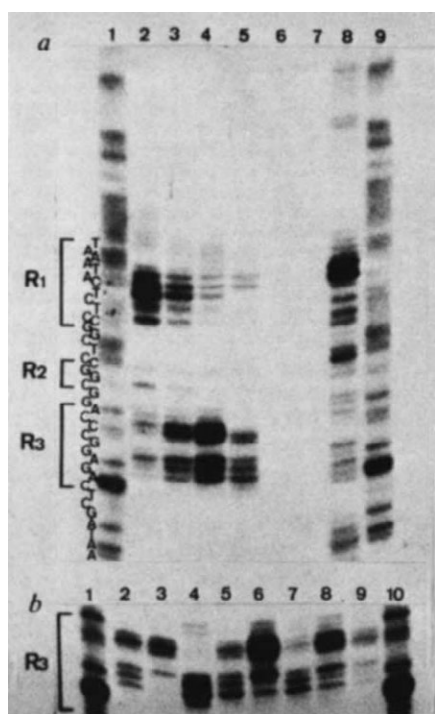
Using a 'run-off' assay<sup>1</sup> we first determined whether RNA polymerase B could initiate accurately *in vitro* on the SV40 early region. A cloned fragment containing the SV40 early sequences *HpaII*-*TaqI* (see Fig. 1 and Fig. 2 legend) was purified and incubated with RNA polymerase B and <sup>32</sup>P-GTP in the presence of a HeLa S100 extract. The newly synthesized RNA was isolated, electrophoresed on a denaturing gel and detected by autoradiography. If initiation of transcription *in vitro* occurred at the major early start sites used *in vivo*, a 'run-off' RNA fragment about 500 base pairs long would be expected (Fig. 1; see below and ref. 13 for details of the assignment of major *in vivo* early cap sites). Transcripts initiating at the major *in vivo* late cap site<sup>14</sup> would appear as short fragments of about 25 base pairs and should therefore not interfere with the assay. Two prominent 'run-off' bands were observed (Fig. 2, lane 2), one (a) about 500 base pairs long and possibly due to initiation at the major *in vivo* early start sites, the other (b) resulting from initiation about 80 base pairs downstream. We determined that



**Fig. 2** Run-off analysis of RNA transcribed *in vitro* from wild-type SV40 early promoter. The 850-base pair *HpaII*-*TaqI* fragment from SV40 was inserted into the *EcoRI* site of pBR322 by blunt-end ligation after DNA polymerase I repair. The SV40 sequences could be excised from the resulting plasmid, pSV3, by *EcoRI* digestion, and were subsequently purified by sucrose gradient centrifugation. The purified fragment was incubated in a HeLa cell-free system and the run-off transcripts analysed as previously described<sup>3,5</sup>, except for the following modifications. (1) The protocol for preparing S100 extract omitted the hypotonic cell washes but included a 30-min incubation of cells at 4°C in hypotonic buffer. Also, the cell homogenate was incubated at 0°C for 30–60 min after being rendered isotonic. (2) The NTP concentration during the RNA synthesis reaction was 200 μM ATP and UTP and 10 μM GTP and CTP. Lane 2, run-off transcripts from the 850-base pair *EcoRI* fragment from pSV3 (containing the SV40 *HpaII*-*TaqI* sequence); lane 3, transcripts from the same *EcoRI* fragment subsequently cut by *Sau3A*I; lane 4, 1 μg ml<sup>-1</sup> α-amanitin was included in the standard assay of transcripts from the pSV3 *EcoRI* fragment; lane 5, exogenous RNA polymerase B was omitted from the standard assay for transcripts from the pSV3 *EcoRI* fragment; lane 6, no DNA was added to the standard assay; lane 7, transcripts of the pSV3 *EcoRI* fragment generated in the absence of S100 extract; lane 9, lane 7 exposed for 1/15th the time; lanes 1 and 8, 5'-end-labelled *HpaII* digests of pBR322. Sizes for these sequenced marker fragments are indicated on the left.

these bands were due to synthesis from the early strand by using an *HpaII*-*Sau* 3AI fragment (Fig. 1a) as template in the *in vitro* system. We found that, as expected for transcripts synthesized from the early strand using this template, the two predominant RNA bands (a and b) obtained were about 30 base pairs shorter (Fig. 2, lane 3). That both bands represented transcription by RNA polymerase B was inferred from their absence when  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin was included in the assay (Fig. 2, lane 4). When RNA polymerase B was omitted, the bands were present but at reduced intensity (Fig. 2, lane 5), indicating that active polymerase B exists in the S100 extract. No bands were detected when DNA was omitted (Fig. 2, lane 6), and in the absence of S100 extract the entire gel lane was blackened. On briefer exposure (lane 9) some bands were visible, but none corresponded to the two predominant start sites found in the presence of S100 extract. Thus, by the run-off assay, the *in vitro* system seems to allow specific initiation of RNA polymerase B at or near the major early cap sites used *in vivo* (a) and at a site about 80 base pairs downstream (b). Handa and co-workers<sup>15</sup>, using a HeLa whole-cell extract, have also detected initiation of transcription in these two regions of SV40.

$S_1$  nuclease mapping gave a more precise determination of the *in vitro* start sites corresponding to run-off band a. The early



**Fig. 3**  $S_1$  nuclease mapping of the 5' ends of RNAs from infected CV1 cells or transcribed *in vitro* in the presence of the HeLa extract. Poly(A)<sup>+</sup> cytoplasmic RNA was isolated from 30-h-infected CV1 cells as described in ref. 13. pSV3 transcripts were synthesized in HeLa extracts as in Fig. 2 (*HpaII*-*TaqI* template) and purified free of DNA as already described<sup>3</sup>. In all cases, 2.5-fold reactions were used for the mapping of *in vitro* transcripts.  $S_1$  nuclease mapping, using a single-stranded *HpaII*-*HinfI* probe  $^{32}\text{P}$ -5'-end labelled at the *HinfI* site (Fig. 1a), essentially was performed as in ref. 13. a, Lanes 2-4, fragments protected by 30-h-infected CV1 cell RNA after 3 h of treatment with 500 (lane 2), 1,000 (lane 3) or 2,000 (lane 4)  $\text{U ml}^{-1}$  of  $S_1$  nuclease; lanes 5-8, fragments protected by *in vitro*-synthesized RNA after 3 h of treatment with 2,500  $\text{U ml}^{-1}$   $S_1$  nuclease. The RNA synthesis reactions contained pSV3 *EcoRI* fragment and the S100 extract in standard conditions (lane 5), pSV3 *EcoRI* fragment and S100 in the presence of  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin (lane 6), S100 in the absence of pSV3 template (lane 7) and pSV3 template in the absence of S100 (lane 8). Lanes 1 and 9 show an A + C sequence ladder of the SV40 *HpaII*-*HinfI* fragment<sup>28</sup>. b, Lane 2, same as lane 4 of a; lanes 3-9, same as lane 5 of a, except that the NTP concentrations during the RNA synthesis reaction were 10  $\mu\text{M}$  GTP and 200  $\mu\text{M}$  CTP, ATP, UTP (lane 3), 10  $\mu\text{M}$  CTP and 200  $\mu\text{M}$  GTP, ATP, UTP (lane 4), 10  $\mu\text{M}$  GTP, CTP and 200  $\mu\text{M}$  ATP, UTP (lane 5), 10  $\mu\text{M}$  ATP and 200  $\mu\text{M}$  UTP, GTP, CTP (lane 6), 10  $\mu\text{M}$  UTP and 200  $\mu\text{M}$  ATP, GTP, CTP (lane 7), 10  $\mu\text{M}$  UTP, ATP and 200  $\mu\text{M}$  GTP, CTP (lane 8), 200  $\mu\text{M}$  GTP, ATP, CTP and UTP (lane 9). Lanes 1 and 10 show an A + C sequence ladder of the SV40 *HpaII*-*HinfI* fragment<sup>28</sup>.

coding strand of an *HpaII*-*HinfI* fragment (Fig. 1a),  $^{32}\text{P}$ -end labelled at the *HinfI* site, was hybridized with either poly(A)<sup>+</sup> cytoplasmic RNA from SV40-infected CV1 cells or RNA transcribed *in vitro* from the *HpaII*-*TaqI* fragment. The hybrids were treated with nuclease  $S_1$ , denatured and electrophoresed on a sequencing gel. Progressively stronger  $S_1$  treatment (Fig. 3a, lanes 2-4) shows that the DNA fragments protected by RNA synthesized *in vivo* has a complex pattern consisting of three components: R1, a cluster of bands mapping between nucleotides 5,189 and 5,197; R2, two bands, at positions 5,175 and 5,177; R3, two clusters of bands mapping at nucleotides 5,164-5,166 and 5,169-5,173. The bands in regions R1 and R2 decrease in intensity with stronger  $S_1$  treatment, while the R3 bands increase. Most of the R1 and R2 bands map within a 27-base hairpin centred at the *BglII* site (Fig. 1c) and we suggest that the transient bands result from  $S_1$  'pauses' in regions of extensive DNA secondary structure (this interpretation has been proposed to explain a similar phenomenon arising during  $S_1$  mapping of conalbumin gene transcripts<sup>3</sup>). With this assumption we conclude that the major *in vivo* early cap sites fall within the region 5,164-5,166 and 5,169-5,173.  $S_1$  mapping does not allow us to assign single-base start sites because  $S_1$  trimming (according to the severity of digestion conditions) can result in 3'-ragged DNA with an overhang of 1-5 nucleotides<sup>16-19</sup>. Nevertheless, our assignment of the major *in vivo* early cap sites agrees with previously published results using reverse transcriptase primer extension mapping<sup>20</sup> or enzymatic cap analysis<sup>13,21</sup>.

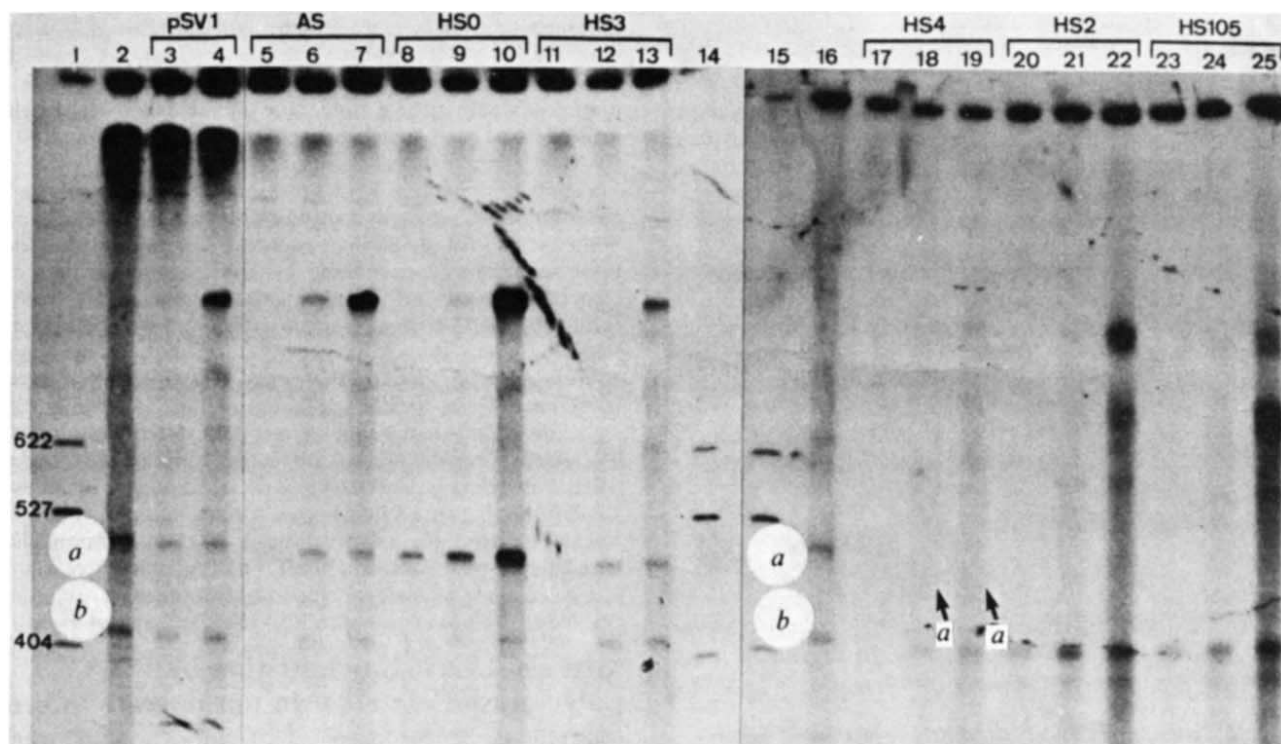
$S_1$  mapping of the *in vitro*-synthesized RNA results in an essentially identical pattern of protected fragments (Fig. 3a, lane 5), suggesting that RNA polymerase B initiates faithfully at the major *in vivo* cap sites, although a low amount of *in vitro* initiation may occur in regions 1 and 2. These  $S_1$ -protected fragments were not observed when  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin was included in the RNA synthesis reaction (Fig. 3a, lane 6), and they require the presence of DNA in the S100 extract (Fig. 3a, lane 7). In the absence of S100 extract (Fig. 3a, lane 8), initiation occurs at multiple sites throughout the fragment.

During these studies we noted a strange phenomenon related to the concentration of nucleotide triphosphates (NTPs) included in the *in vitro* RNA synthesis reaction. As described above, the major start sites *in vitro* consist of two clusters of bands (R3); however, this is not true if the concentration of purines in the synthesis reaction greatly exceeds that of pyrimidines or vice versa. For example, when GTP (or ATP) is present at 10  $\mu\text{M}$  and the other three NTPs at 200  $\mu\text{M}$ , the upper cluster of bands in R3 is much more prominent (Fig. 3b, lanes 3, 6). In contrast, when CTP (or UTP) is included at 10  $\mu\text{M}$  and the other three NTPs at 200  $\mu\text{M}$ , the lower cluster predominates (Fig. 3b, lanes 4, 7). Only if the total purine concentration approximates the total pyrimidine are both clusters of bands observed in a ratio similar to that found with *in vivo* RNA (Fig. 3b, compare lane 2 with lanes 5, 8 and 9). Although we cannot explain this phenomenon, it does suggest that our  $S_1$  mapping detects true RNA polymerase start regions, as it is unlikely that the NTP concentration affects anything other than the *in vitro* RNA synthesis reaction. Also, any *in vitro* study of mutant promoter regions must consider this NTP concentration effect.

$S_1$  mapping of the RNA represented by run-off fragment b using the *EcoRII* G fragment (Fig. 1a) (data not shown) shows that the 5' end of the b transcript maps around nucleotide 5,080. Bands were not detected at this site in  $S_1$  mapping of *in vivo* wild-type RNA even on heavily overloaded gels (although a band at 5,085 can sometimes be observed; data not shown). A 5'-GATAAA-3' sequence occurs 15 base pairs upstream from position 5,080, and a hairpin spans nucleotides 5,100-5,120 (centred ~30 base pairs upstream from the start site). We do not know whether either of these sequence elements corresponds to an initiation signal recognized *in vitro* but not *in vivo*.

Thus, we have shown that RNA polymerase B initiation on the SV40 early region occurs *in vitro* at a limited number of sites,





**Fig. 4** Run-off analysis of RNA transcribed *in vitro* from deletion mutants in the promoter region. The construction of the wild-type plasmid pSV1 and the derivation of plasmids bearing deletions in the promoter region are described in ref. 13. Each of these recombinants was cut by *TaqI* and *HpaII* to provide an SV40 early region truncated template. The fragments were purified on sucrose gradients and then incubated for RNA synthesis in the presence of the HeLa S100 extract. The incubation conditions and subsequent analysis of transcripts were exactly as described in Fig. 2 legend. The RNA synthesis reactions contained for lanes 2 and 16:  $4 \mu\text{g ml}^{-1}$  of the 850-base pair *EcoRI* fragment from pSV3; for lanes 3 and 4, 4 and  $8 \mu\text{g ml}^{-1}$ , respectively, of the *HpaII-TaqI* fragment from pSV1; for the remaining lanes: 2, 4 or  $8 \mu\text{g ml}^{-1}$  of the *HpaII-TaqI* fragment from AS (lanes 5–7), HS0 (lanes 8–10), HS3 (lanes 11–13), HS4 (lanes 17–19), HS2 (lanes 20–22) or HS105 (lanes 23–25). The arrows in lanes 18 and 19 indicate the position of the band *a* transcripts visible on the original autoradiogram. All samples were derived from the same experiment although they were electrophoresed on two gels (lanes 1–14 and 15–25). The markers in lanes 1, 14 and 15 were  $^{32}\text{P}$ -5'-end labelled *HpaII* fragments from pBR322.

one prominent set corresponding to the major cap sites used *in vivo*. Transcription seems rather inefficient, however, being about 20% that of the Ad-2 major late transcription unit in the same conditions (data not shown). Nevertheless, the fidelity of initiation with the HeLa cell-free system allows us to approach the question of whether mutations in the promoter region affect *in vitro* and *in vivo* initiation of transcription identically.

### Some *in vivo* transcription signals are recognized *in vitro* by RNA polymerase B

The accompanying article<sup>13</sup> describes the construction of a series of deletion mutants spanning the promoter region for the SV40 early genes. Deletions were of three types (Fig. 1b): (1) those removing sequences between the TATA box and the major *in vivo* cap sites (for example, AS), (2) those which start at position 5,162, delete the major cap sites and extend upstream towards and eventually beyond the TATA box (the HS series), and (3) those starting from the *HpaII* site at position 282 and extending downstream towards the TATA box (the PS series). RNA synthesized *in vitro* by the first two classes of mutants indicated that one function of the TATA sequence was to fix initiation of transcription within a narrow region<sup>13</sup>. In its presence, RNA polymerase B starts transcription about 30 base pairs downstream; in its absence, the enzyme initiates at multiple sites over a broad region. We now assess the effect of these promoter region deletions on *in vitro* RNA synthesis.

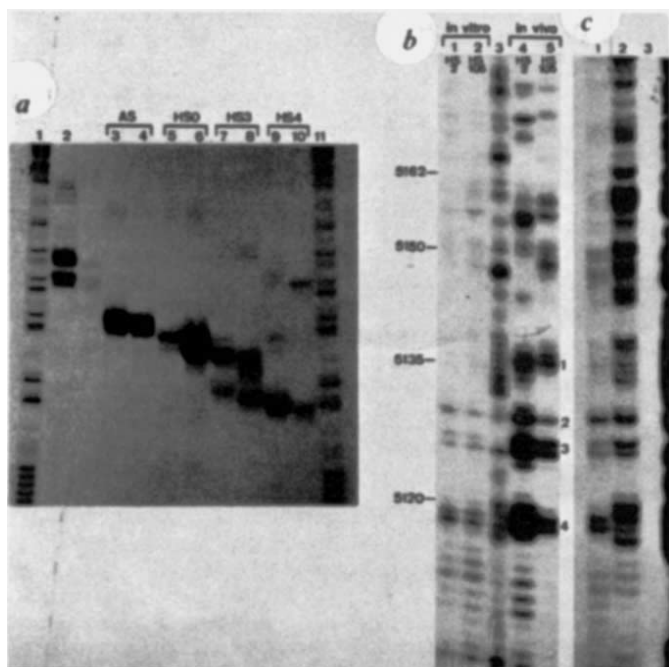
*HpaII-TaqI* fragments purified from the wild-type 'parental' plasmid pSV1 and from the various mutants, were transcribed *in vitro* and the run-off RNA transcripts analysed as described above (Fig. 2) for wild-type plasmid pSV3. (Note that pSV3 contains the SV40 early sequences *HpaII-TaqI* inserted at the *EcoRI* site of pBR322, whereas pSV1 contains the *HpaII-BamHI* sequences inserted at the same site.) Incubation of the pSV1 fragment in the cell-free system resulted in the expected

500 (*a*) and 420 (*b*) base pair run-off bands (Fig. 4, lanes 3, 4). Some bands not observed with the other wild-type plasmid—pSV3—were also evident (compare lanes 2 and 4), but these are not  $\alpha$ -amanitin sensitive (data not shown) and probably correspond to full-length transcripts of contaminating DNA fragments [*HpaII-TaqI* sequences are easier to purify from pSV3 than from pSV1 or the various deletion mutants, especially when large deletions are involved (see Figs 2 and 4 legends for details)].

Mutants AS, HS0, HS3 and HS4 all retain the TATA box, although only AS retains the major *in vivo* cap sites (Fig. 1b). Run-off transcripts *a* and *b* are clearly produced when AS (Fig. 5, lanes 5–7), HS0 (lanes 8–10) and HS3 (lanes 11–13) fragments are incubated in the cell-free system. With HS4 (lanes 17–19), the *a* and *b* bands are both detectable, at least on the original autoradiogram, but *a* is much reduced. Closer inspection of the data reveals that with deletions which do not involve the TATA box (1) the size of band *a* decreases progressively with larger deletions, and (2) the ratio of *a* to *b* is different with the various mutants. Densitometric scanning of the autoradiograms indicates the following *a/b* ratios: pSV1 or pSV3, 1.0; AS, 2.5; HS0, 6.5; HS3, 1.0; HS4, 0.25. This effect has been observed in several experiments using different DNA preparations and varying NTP concentrations.

Mutants HS2 and HS105, which lack both the TATA box and the major *in vivo* cap sites (Fig. 1b), do not express an *a* transcript (lanes 20–22 and 23–25), although band *b* is readily apparent (a result obtained with several different DNA preparations and in all NTP conditions studied).

Using S<sub>1</sub> nuclease mapping to compare the effect of promoter region deletions on initiation *in vitro* and *in vivo* (Fig. 5a), we found that for those mutants which retain the TATA box (AS, HS0, HS3 and HS4), the major *in vivo* start sites are mimicked quite faithfully *in vitro*. In both environments, RNA polymerase B initiates about 30 base pairs downstream from the TATA box,



**Fig. 5**  $S_1$  nuclease mapping of 5' ends of RNAs synthesized *in vivo* or *in vitro* from mutant plasmids. Poly(A)<sup>+</sup> cytoplasmic RNA was isolated from rat 3T3 cells transformed by the various mutants as described in ref. 13.  $S_1$  mapping was as in Fig. 3 except that *HpaII*-*HinfI* probes specific for each mutant were used. *a*, The cellular RNAs were from cell lines transformed by pSV1 (lane 2), AS (lane 3), HS0 (lane 5), HS3 (lane 7) or HS4 (lane 9). The *in vitro* RNA synthesis reactions contained *HpaII*-*TaqI* fragments from plasmids AS (lane 4), HS0 (lane 6), HS3 (lane 8) and HS4 (lane 10). Lanes 1 and 11 show an A+C sequence ladder<sup>28</sup> from the wild-type *HpaII*-*HinfI* fragment. *b*, The cellular RNAs were from cell lines transformed by HS2 (lane 4) and HS105 (lane 5). The *in vitro* RNA synthesis reactions contained *HpaII*-*HhaI*-*TaqI* fragments from plasmids HS2 (lane 1) and HS105 (lane 2) (for mutants with long deletions, *HhaI* digestion was performed to facilitate separation of the SV40 fragment from contaminating pBR322 sequences). Lane 3 shows an A+C sequence ladder<sup>28</sup> from the wild-type *HpaII*-*HinfI* fragment. The numbers to the left refer to nucleotide sequence number as in Fig. 1a; the numbers to the right correspond to band numbers in Fig. 3b of ref. 13. *c*, The *in vitro* RNA synthesis reactions contained *HpaII*-*HhaI*-*TaqI* fragments from HS105 in standard conditions (lane 1) in the absence of S100 (lane 2) or in the presence of  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin (lane 3). For lane 2, only 1/8 of the  $S_1$ -protected material was loaded.

regardless of the exact downstream sequences (ref. 13 gives a precise map of the *in vivo* start sites for these mutants). There are some differences in the RNAs synthesized *in vivo* and *in vitro*: for HS0, the ratio of the two starts differs in the two environments; for HS3 the *in vitro* start sites appear to extend further downstream than those of *in vivo* RNA. The experiments depicted in Fig. 5 give no indication of the efficiencies of transcription, as DNA probes of different specific activity were used for each mutant. However,  $S_1$ -mapping of the mutant transcripts using a single wild-type probe (data not shown) shows that the level of initiation just downstream from the early region TATA box varied for different mutants, in the order HS0 > AS > HS3  $\approx$  PSV1 > HS4.

When the TATA box is deleted, early gene expression can still be detected *in vivo*, but transcription begins at multiple sites over a broad area<sup>13</sup>, and  $S_1$  mapping experiments (Fig. 5b, c) demonstrate that this phenomenon also occurs *in vitro*. (This multiplicity of initiation sites probably explains why no distinct bands can be detected by the run-off assay; see above.) For a given mutant, the pattern of fragments protected by *in vitro*-synthesized RNA closely resembles that of fragments protected by *in vivo* transcripts (compare lanes 1 and 4 or lanes 2 and 5; see also Fig. 3b in ref. 13). Most bands are observed with both *in vivo* and *in vitro* RNAs, although relative band intensities differ, sometimes markedly. Using a single wild-type probe we also determined that the *in vitro* transcription of HS2 is not much less than that of the wild-type plasmid pSV1 (data not shown),

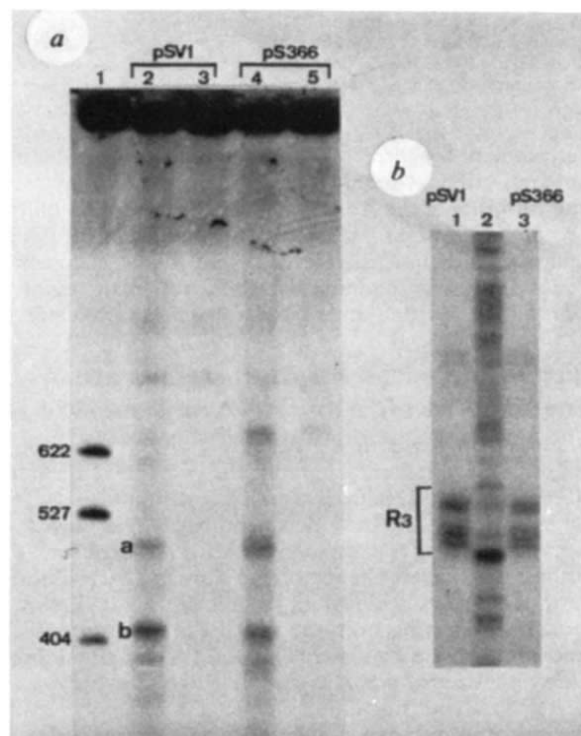
although the new set of cap sites generated in HS2 was not detected for pSV1, indicating that, as *in vivo*<sup>13</sup>, the wild-type TATA sequence may 'mask' its functional substitutes. We do not know whether these 'substitute' elements are preferentially located in the early promoter region or scattered throughout the SV40 genome.

Figure 5c shows that RNA polymerase B alone selects specific initiation sites, but yields a very heterogeneous pattern (lane 2). Some factor(s) of the S100 extract restricts initiation to a few of these sites (lane 1), producing a similar pattern to that seen *in vivo*. It also requires RNA polymerase B, being abolished by the inclusion of  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin in the RNA synthesis reaction (lane 3).

We therefore conclude that some of the initiation 'signals' recognized *in vivo* by RNA polymerase B are also recognized *in vitro*. Initiation occurs about 30 base pairs downstream from the TATA box when it is present, but at many sites over a broad area when it is deleted. The 'fixing' function probably includes the TATA box because initiation  $\sim 30$  base pairs downstream is already reduced in HS4, which has a deletion end point just 2 base pairs downstream from the TATTTAT sequence (Fig. 1c), but we cannot exclude the possibility that other sequences, deleted in HS2 but present in HS4 (Fig. 1c), are involved.

### 'Signals' controlling initiation by RNA polymerase B *in vivo* do not operate *in vitro*

The third class of deletion mutants described in the accompanying article<sup>13</sup>, the PS series, retains the major wild-type cap sites as well as a normal TATA box, but lacks sequences further upstream. Analysis of the *in vivo* expression of PS mutants



**Fig. 6** Run-off analysis and  $S_1$  nuclease mapping of RNA synthesized *in vitro* from the deletion mutant PS366. The construction of PS366 is described in ref. 13. *a*, Run-off analysis was as in Fig. 2. Lane 2, the *HpaII*-*TaqI* fragment from pSV1 was incubated with S100 extract in standard conditions. Lane 3, the same fragment incubated with S100 in the presence of  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin. Lane 4, the *HpaII*-*HhaI*-*TaqI* fragment from PS366 was incubated with S100 in standard conditions. Lane 5, the same fragment was incubated with S100 in the presence of  $1 \mu\text{g ml}^{-1}$   $\alpha$ -amanitin. Lane 1 shows a  $^{32}\text{P}$ -5'-end-labelled *HpaII* digest of pBR322. *b*,  $S_1$  nuclease mapping was as in Fig. 5. Lane 1, fragments protected by RNA synthesized *in vitro* from the pSV1 template; lane 2, an A+C sequence ladder<sup>28</sup> of the wild-type *HpaII*-*HinfI* fragment; lane 3, fragments protected by RNA synthesized *in vitro* from the PS366 template. R3 refers to the region delineated in Fig. 3a.



indicates that sequences mapping in the region of the two 72-base pair repeats I and II (Fig. 1b) markedly affect early gene expression. For example, PS366, which lacks nucleotides 98–282 and thus more than one complete copy of the 72-base pair repeat (Fig. 1b), does not detectably express the early genes during short-term infection of CV1 cells, and transforms rat 3T3 fibroblasts much less efficiently than a wild-type plasmid. However, Fig. 6a demonstrates that, by the run-off assay, *in vitro* transcription of PS366 is not similarly inefficient. The *a* transcript is readily apparent when an *HpaII*–*TaqI* fragment from PS366 is transcribed in the HeLa cell-free system (lane 4). Furthermore, S<sub>1</sub> nuclease mapping reveals that RNA polymerase B initiates the *a* transcript at the same sites on PS366 as on the wild-type pSV1 (Fig. 6b, lanes 1, 3). Thus, those sequence elements in the region of the 72-base pair repeats that exert a crucial influence on *in vivo* expression have no apparent effect in an *in vitro* environment.

### ***In vitro* and *in vivo* promoter regions for RNA polymerase B**

The three DNA sequence elements implicated in the control of initiation of transcription by RNA polymerase B are the TATA box region<sup>3,5,10,11,13</sup>, the cap site region<sup>5,11</sup> and a region upstream from the TATA box (refs 11–13, and personal communications from R. Grosschedl and M. L. Birnstiel, and from R. Kamen).

The present report and the accompanying one suggest that, both *in vivo* and *in vitro*, the early gene TATA box (or closely adjacent sequences) functions to limit initiation to a narrow region. In its presence, RNA polymerase B starts transcription about 30 base pairs downstream, and a similar function has been noted for other genes<sup>5,11,13,22</sup>. Whatever actually measures the TATA box–cap site distance operates quite accurately *in vitro*, even in a heterologous system, for the newly selected cap sites on mutant templates are similar to those used in mammalian cells. The finding that on deletion of the early gene TATA box RNA polymerase initiates at multiple sites over a broad region both *in vitro* and *in vivo* is in keeping with a report that, in *Xenopus* oocytes, initiation occurs at multiple sites on an injected sea urchin histone gene lacking its normal TATA sequence<sup>11</sup>. The multiple initiation sites we observe are nearly identical to those found *in vivo* (Fig. 5b), and this specific pattern is dependent on the presence of some factor(s) present in the cell-free extract (Fig. 5c), suggesting that the factor(s) which recognize the TATA box could also recognize its functional substitutes. Deletions or single-point mutations involving the TATA box have been reported to abolish, or drastically reduce, *in vitro* transcription<sup>5,10</sup>, but such studies relied mainly on the run-off assay, and we also were unable to detect transcripts from the TATA-minus mutants using this assay. Alternatively, it is possible that some genes, like the SV40 early region, have sequences which can substitute functionally for the TATA box, while other genes either do not have them or have sequences which function much less efficiently than the TATA sequence.

We have observed that in mutants retaining the TATA box, the newly selected cap sites seem to be used with varying efficiencies *in vitro*, and this has been attributed<sup>5,11</sup> to control by sequences around the cap site. We cannot decide whether the crucial sequences are those at the cap site and/or those between the TATA box and start site. Clearly, the mere presence of purines around the initiating nucleotide is not the important factor, because initiation on both HS0 and HS4 occurs in a region rich in purines. (*In vivo* transcription of many genes by RNA polymerase B starts with an A within a pyrimidine-rich region<sup>5,9</sup>.) Nucleotides between the TATA box and the start sites might well influence the efficiency of cap site utilization, especially for the SV40 early genes, where a 27-base hairpin normally separates the TATA and the cap site (Fig. 1c). It is possible that AS and HS0 are transcribed so efficiently merely because this hairpin is disrupted, and that HS4 is transcribed much less efficiently because its longer deletion begins to affect the TATA box function. Interestingly, similar 'up'-mutations have been noted *in vivo*. Rat cells transformed by plasmids

bearing deletions 10–15 base pairs downstream from the early gene TATA box produce more early mRNA than cells transformed by wild-type plasmid (see ref. 22 and the accompanying article). Gluzman *et al.*<sup>22</sup> have suggested that this overproduction of early mRNA might be due to a failure in autoregulation by T antigen. In light of the *in vitro* results, it might, instead or in addition, be caused by a more efficient utilization of the newly selected cap sites.

The far upstream sequences<sup>13</sup>, which are required *in vivo* but not *in vitro*, might bind RNA polymerase directly (analogous in some respects to the prokaryotic 'recognition site'<sup>23</sup>), might dictate chromatin structure directly<sup>24</sup> or indirectly, or might bind activator proteins. We have shown that the mutant plasmid PS366, which lacks certain sequences upstream from the TATA box and is expressed inefficiently *in vivo*, is transcribed quite readily in the HeLa cell-free extract. The level of initiation approximates that for wild-type plasmid and the start sites are those normally used. This observation tends to argue against a model whereby the upstream sequences exert their influence simply by binding polymerase. However, the efficiency of transcription *in vitro* is much lower than *in vivo*<sup>1–5</sup>, and it is possible that the mode of polymerase binding differs in the two environments. It may be that the chromatin structure in the promoter region brings into close vicinity the TATA box and the upstream sequences and that accurate reconstruction of chromatin structure might be important in understanding the role of the upstream sequences. We do not know the state of the DNA after incubation in the S100 extract, but it is unlikely to be packaged in chromatin as it is *in vivo*, for we have observed that during the standard RNA synthesis reaction supercoiled DNA templates are converted to relaxed covalently closed circles (unpublished results), whereas if nucleosomes had been formed, supercoiled DNA should have been recovered<sup>23,24</sup>. Certainly, activator proteins could also be missing from the cell-free system. Clues as to how the upstream sequences control initiation *in vivo* and why they do not exert a similar influence *in vitro* await further experimentation.

We thank Drs C. Benoist, B. Davison and C. Kédinger for the gift of materials and for helpful advice, Mr J. M. Garnier for technical assistance, and Ms B. Chambon, C. Kutschis and Mr B. Boulay for help in preparing the manuscript. This work was supported by grants from the INSERM (ATP 72.79.104, CL 801029 and CL 801033), the CNRS (ATP 'Biologie Moléculaire de Gène'), the Fondation Française pour la Recherche Médicale and from the Fondation S. and C. Del Duca.

Received 1 October 1980; accepted 22 January 1981.

- Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18**, 469–484 (1979).
- Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gelfand, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855–3859 (1980).
- Waslylyk, B., Kédinger, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367–373 (1980).
- Luse, D. S. & Roeder, R. G. *Cell* **20**, 691–699 (1980).
- Corden, J. *et al. Science* **209**, 1406–1414 (1980).
- Gannon, F. *et al. Nature* **278**, 428–434 (1979).
- Proudfoot, N. J. *Nature* **279**, 376 (1979).
- Flavell, R. A. *Nature* **285**, 356–357 (1980).
- Breathnach, R. & Chambon, P. A. *Rev. Biochem.* (in the press).
- Waslylyk, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7024–7028 (1980).
- Grosschedl, R. & Birnstiel, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432–1436 (1980).
- Benoist, C. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3865–3869 (1980).
- Benoist, C. & Chambon, P. *Nature* **290**, 304–310 (1981).
- Ghosh, P., Reddy, V., Swinscoe, J., Lebowitz, P. & Weissman, S. *J. molec. Biol.* **126**, 813–846 (1979).
- Handa, H., Kaufman, R. J., Manley, J., Gelfand, M. & Sharp, P. A. *J. biol. Chem.* (in the press).
- Hentschel, C., Irminger, J. C., Bucher, P. & Birnstiel, M. L. *Nature* **285**, 147–151 (1980).
- Weaver, R. F. & Weissman, C. *Nucleic Acids Res.* **7**, 1175–1195 (1979).
- Sollner-Webb, B. & Reeder, R. *Cell* **18**, 485–499 (1979).
- Tsujimoto, Y. & Suzuki, Y. *Cell* **16**, 425–436 (1979).
- Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. *J. Virol.* **30**, 279–296 (1979).
- Haegeman, G. & Fiers, W. *J. Virol.* **35**, 955–961 (1980).
- Gluzman, Y., Sambrook, J. F. & Frisque, R. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3898–3902 (1980).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).
- Waslylyk, B., Oudet, P. & Chambon, P. *Nucleic Acids Res.* **7**, 705–713 (1979).
- Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847 (1975).
- Birkenmeier, E. H., Radonovich, M. F., Shani, M. & Salzman, N. P. *Cell* **11**, 495–504 (1977).
- Van Heuverswyn, H. & Fiers, W. *Eur. J. Biochem.* **100**, 51–60 (1979).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–580 (1980).

## LETTERS

Pulsar and diffuse contributions to observed galactic  $\gamma$  radiation

A. K. Harding &amp; F. W. Stecker

Laboratory for High Energy Astrophysics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

Satellite data on the energy spectrum of galactic  $\gamma$ -radiation<sup>1,2</sup>, shows that such radiation has a multicomponent nature<sup>3-7</sup>. Here we use a calculation of the pulsar  $\gamma$ -ray emission spectrum together with a statistical analysis of recent data on 328 known pulsars to make a new determination of the pulsar contribution to galactic  $\gamma$ -ray emission. We then re-examine the contributions from diffuse interstellar cosmic-ray induced production mechanisms to the total emission and conclude that pulsars may account for a significant fraction of galactic  $\gamma$ -ray emission.

Models for the production of high-energy  $\gamma$  rays in the Galaxy have dealt primarily with diffuse emission processes, those which involve interactions between cosmic rays and interstellar matter. The major contributions are thought to come from electron bremsstrahlung and the decay of neutral pions produced in high energy proton collisions, with a small additional flux from Compton interactions. In determining the relative contributions of these processes from the shape of the observed  $\gamma$ -ray spectrum, the contribution of true point sources (as opposed to enhancements in the spatial distribution from diffuse processes) has been neglected, largely because the point source spectrum has been undetermined. At present, pulsars can give the best information on a true point source component of the galactic  $\gamma$ -ray flux.

We report here on the first calculation of a  $\gamma$ -ray production spectrum from pulsars in the Galaxy and on the implications of this point source contribution to the general interpretation of the observed galactic  $\gamma$ -ray spectrum. The details of this calculation, as well as a determination of the longitude and latitude profiles of pulsar  $\gamma$  rays will be described elsewhere<sup>8</sup>. There have been previous estimates of the pulsar contribution to the observed galactic  $\gamma$ -ray flux<sup>9-11</sup>, using only 51 pulsars and very general assumptions about their  $\gamma$ -ray luminosities. There are now 328 known radio pulsars, all having independent distance determinations through their dispersion measures. The local density and galactic distribution can therefore be much better determined, assuming a value for the mean free electron density in the galactic plane. Pulsed  $\gamma$  rays have been observed from the Crab (PSR0531+21) and Vela (PSR0833-45) pulsars by the SAS 2 satellite<sup>12,13</sup>, and the COS B satellite<sup>14</sup>. The energy spectra of these sources have recently been determined in the range 50 MeV–2 GeV and they can both be fit with steep power laws of similar slope around  $-2$  (ref. 15). Although these two young sources represent only a small fraction of the known pulsars, their emission testifies to the fact that some, and possibly all pulsars are  $\gamma$ -ray emitters at some stage of their evolution. Though these others may be individually below the present limits of detectability, the contribution of their  $\gamma$  rays to the observed galactic emission could be significant, as has been previously suggested<sup>9</sup>.

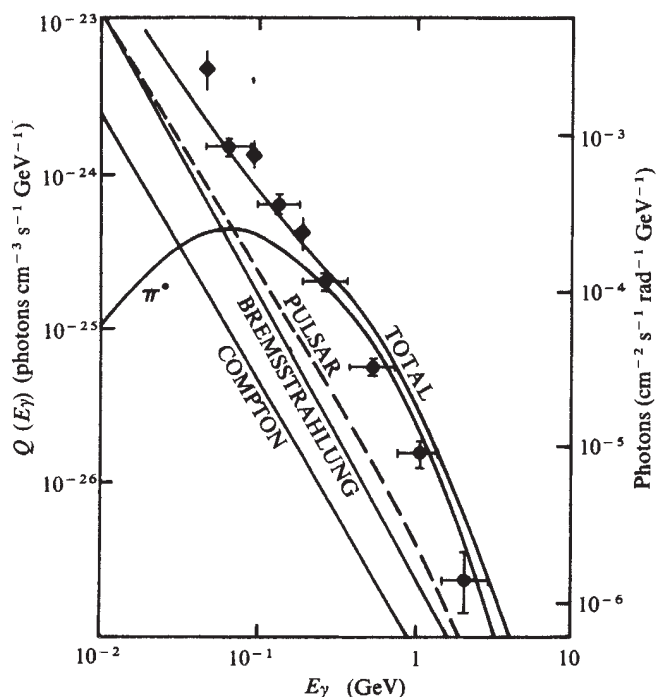
We assume that the  $\gamma$ -ray emission above 100 MeV is produced by curvature radiation from primary particles which have been accelerated to high energies in the electric fields of the pulsar's magnetosphere. Some of these  $\gamma$  rays are converted to electron–positron pairs in the strong pulsar magnetic field. The electric field induced perpendicular to the magnetic field by the star's rotation also contributes to the pair production rate and has a large effect on the optical depths in short period pulsars<sup>16</sup>. The shape of the numerically calculated spectra in this model

depend on the initial energy of the primary particles and on the pulsar magnetic field strength and period<sup>17</sup>. For an initial energy of  $1.5 \times 10^4$  GeV and a field strength of  $10^{12}$  G, the resulting calculated spectrum for a pulsar period of 0.033 s reproduces the shape of the observed Crab pulsar spectrum. The observed 100-MeV flux from the Crab pulsar at 2 kpc is used to determine the normalization factor ( $\sim 10^3$ ) which is the ratio of the primary particle density above the polar cap to the corotation charge density. Using this same normalization factor, primary particle energy, and magnetic field strength, the calculated spectrum for a period of 0.089 s fits the observed Vela pulsar spectrum quite well. It may be reasonable to assume, then, that these values of the parameters determining the spectrum are valid for all pulsar periods.

An expression for the  $\gamma$ -ray luminosity can be derived by integrating the spectra calculated for different periods,  $P$ , and surface magnetic field strengths,  $B_{12}$  (in units of  $10^{12}$  G). The best fit for the dependence of luminosity on these quantities gives:

$$L_{\gamma}(>100 \text{ MeV}) = 1.2 \times 10^{35} B_{12}^{0.95} P^{-1.7} \text{ photons s}^{-1} \quad (1)$$

The  $\gamma$ -ray luminosities and fluxes for the known radio pulsars have been calculated from equation (1) and compared with the upper limits obtained by SAS 2 (ref. 18) and the sources in the COS B catalogue<sup>19</sup>. Eight pulsars have predicted fluxes above the point source detectability limit of COS B but most are towards the galactic centre where the diffuse background is high. Only two lie in the longitude range  $90^\circ < l < 270^\circ$  where the search is considered complete down to  $1.3 \times 10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$ ; however, they both lie just outside the latitude range of the search. Four pulsars have predicted fluxes above their SAS 2 upper limits, but exceed them by no more than the estimated uncertainties in either the measurements or the model.



**Fig. 1** Differential production spectra for pulsars and the major diffuse processes. The  $\pi^0$ , bremsstrahlung and Compton spectra are from refs 6, 26 and the total spectrum includes the pulsar contribution. SAS 2 and COS B data points for the longitude range  $355^\circ < l < 15^\circ$  are also shown ( $\diamond$ , ref. 2;  $\bullet$ , ref. 1). The left- and right-hand scales refer to the theoretical production rates at 10 kpc from the galactic centre and the observed  $\gamma$ -ray flux, respectively.



In calculating these pulsar spectra and luminosities, it has been assumed that the magnetic field is a pure dipole and that the angle between the dipole axis and spin axis is  $90^\circ$ , because we are concerned here with the collective properties of the pulsar population, and a random distribution of angles would give an average of  $90^\circ$ . More details of the model are given in ref. 17.

If all pulsars produce pulsed  $\gamma$  rays according to this model, then the local differential production spectrum will be,

$$q_{\text{PSR}}(E_\gamma) = \frac{D_\odot}{2h} \int_0^\infty I_\gamma(P, E_\gamma) \rho(P) dP$$

where  $D_\odot$  is the local surface density of pulsars which are beamed in our direction,  $h$  is their scale height above the plane,  $\rho(P)$  is their period distribution and  $I_\gamma(P, E_\gamma)$  is the  $\gamma$ -ray spectrum of an individual pulsar with period  $P$ .  $I_\gamma(P, E_\gamma)$  is also directly proportional to surface magnetic field strength, so that  $q_{\text{PSR}}$  will be proportional to the average pulsar magnetic field strength. We take an average field of  $10^{12}$  G, which is consistent with measured pulsar spindown rates. The derived local density and scale height depend on a determination of  $\langle n_e \rangle$ , the mean electron density in the galactic plane, which, through dispersion measure observations, fixes the pulsar distance scale. Unfortunately, the galactic distribution of  $n_e$  is not well known and is difficult to determine. There are several indications that  $n_e$  is not uniform throughout the galactic plane and that it does not have a simple exponential  $Z$  dependence<sup>20,22</sup>, both of which have been previously assumed in determining pulsar densities. However, in view of the difficulties entailed in a detailed spatial modeling of  $n_e$ , given the present observational data, we adopt the simple constant density model of Taylor and Manchester<sup>23</sup> which is sufficient for estimating the mean contribution from pulsars to the galactic  $\gamma$ -radiation. For  $\langle n_e \rangle = 0.03 \text{ cm}^{-3}$ , a value consistent with HI absorption distances to pulsars in local regions of the Galaxy<sup>20</sup>, a statistical analysis of data on the 328 known pulsars gives  $D_\odot = 480 \text{ pulsars kpc}^{-2}$  and  $h_p \approx 325 \text{ pc}$ . The resulting  $\gamma$ -ray production spectrum per  $\text{cm}^3$  is shown in Fig. 1 along with the production spectra for the major diffuse processes calculated<sup>6</sup>, assuming a local interstellar hydrogen density of  $1 \text{ atom cm}^{-3}$ . This spectrum gives volume-averaged integral pulsar  $\gamma$ -ray production rates at 10 kpc of  $2.2 \times 10^{-26}$  photons  $\text{cm}^{-3} \text{ s}^{-1}$  above 100 MeV and  $4.6 \times 10^{-26}$  photons  $\text{cm}^{-3} \text{ s}^{-1}$  above 35 MeV. Although the calculated pulsar spectrum has been extrapolated below 100 MeV, the model does not accurately predict the low energy spectrum because the synchrotron radiation from secondary particles, which may extend to the  $\gamma$ -ray region, has not been included. This contribution would tend to steepen the spectrum and increase the estimated flux level below around 100 MeV.

The diffuse production spectra in Fig. 1 have widely differing uncertainties. The pion production spectrum is least uncertain, because it primarily involves cosmic ray protons with energies in the range 1–10 GeV where uncertainties in determining galactic fluxes owing to solar modulation effects produce small uncertainties ( $\approx 15\%$ ) in the total production rate<sup>24</sup>. Even the differential  $\gamma$ -ray source spectrum for  $E_\gamma \leq 10 \text{ GeV}$  is insensitive to the kinematical model chosen, as shown by the excellent agreement between calculations given in refs 24–28.

The Compton production mechanism produces only a small expected contribution in our region of the Galaxy, however, this process could be significant near the galactic centre<sup>6</sup>. As this component is small over most of the galaxy and various theoretical estimates<sup>29,30</sup> are in reasonable agreement with Fig. 1, we conclude that little error in determining the total galactic flux is introduced by the Compton process.

The situation is quite different in the case of electron bremsstrahlung. The curve shown in Fig. 1 for this component was determined using a conservative 'standard' cosmic ray electron spectrum<sup>31</sup>. However, the electron spectrum cannot be easily determined from cosmic-ray data due to large solar modulation effects. Theorists have thus had much leeway in invoking large electron fluxes to account for discrepancies in both the spectrum and intensity of the observed galactic  $\gamma$  rays with 'standard'

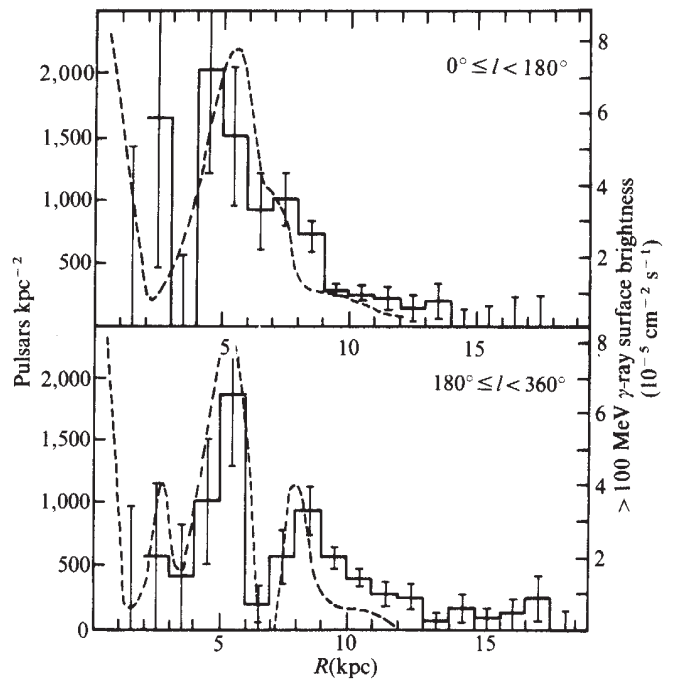


Fig. 2 Surface density of pulsars (solid line)<sup>8</sup> and  $\gamma$ -ray surface emissivity (dashed line)<sup>6</sup> plotted against galacto-centric radius  $R$  for northern and southern halves of the Galaxy.

predictions<sup>32,33</sup>. Indeed, spectral data have been used to speculate on drastic fluctuations in both electron and proton fluxes in the Galaxy<sup>34</sup>, although we believe that inherent systematic errors as well as statistical errors in the spectral data make such a radical analysis unwarranted.

The present calculated pulsar contribution has clear implications for analysing the total galactic spectrum. Pulsars contribute a soft power law flux component similar to that expected from electron bremsstrahlung. In addition, the pulsar component as estimated in Fig. 1 is as large as the 'standard' bremsstrahlung component.

We also expect that the pulsar contribution to the total galactic flux should be roughly independent of galactic longitude except in the outer Galaxy. We have used the longitude and distance data for the full sample of observed pulsars, together with corrections for selection effects, to determine the radial distributions of pulsars in both the northern ( $0^\circ < l < 180^\circ$ ) and southern ( $180^\circ < l < 360^\circ$ ) galactic longitudes. The results for the inner Galaxy ( $R \leq 8 \text{ kpc}$ ) are strikingly similar to the distributions of total  $\gamma$ -ray emissivity determined previously<sup>6</sup>. Thus, the ratio of pulsar to total emissivity (the fraction of the flux from the pulsar component) should be roughly constant with longitude, for longitudes within  $60^\circ$  of the galactic centre. This fraction is calculated to be 15–20% above 100 MeV (as derived from both Fig. 1 and a comparison of the longitude distribution of pulsar and observed  $\gamma$ -ray emission<sup>8</sup> using the data in Fig. 2). It can easily be of the order of 50% at lower energies (see Fig. 1), removing the need to invoke large low-energy electron fluxes to explain the shape of the observed  $\gamma$ -ray spectrum.

The latitude distribution of pulsar  $\gamma$ -ray emission will be narrow but will extend somewhat beyond  $b = 10^\circ$  (ref. 8). Pulsars may therefore be able to account for some of the flux observed<sup>35,36</sup> at galactic latitudes above  $10^\circ$ . According to the model, the younger pulsars are the strongest  $\gamma$ -emitters. They are also more confined to the galactic plane<sup>37</sup>. Thus, the pulsars at high  $b$  have a higher mean age and are weaker. This would be consistent with the lack of observed point sources at high  $b$ .

We conclude therefore that the integrated flux from distant  $\gamma$ -ray emitting pulsars may provide an important fraction of the low- to medium-energy galactic  $\gamma$ -ray flux (although a smaller fraction of the flux above 100 MeV), and the inclusion of the

pulsar component in analysing the galactic  $\gamma$ -ray spectrum supports a coherent integrated relationship with galactic cosmic-ray theory as discussed elsewhere<sup>38-42</sup>.

Received 16 October 1980; accepted 17 February 1981.

1. Paul, J. A. *et al.* *Astr. Astrophys.* **63**, L31-L33 (1978).
2. Hartman, R. C. *et al.* *Astrophys. J.* **230**, 597-606 (1979).
3. Stecker, F. W. *Cosmic Gamma Rays* (Mono, Baltimore, 1971).
4. Stecker, F. W., Puget, J. L., Strong, A. W. & Bredekamp, J. H. *Astrophys. J. Lett.* **188**, L59-L61 (1974).
5. Fichtel, C. E. *et al.* *Astrophys. J.* **208**, 211-219 (1976).
6. Stecker, F. W. *Astrophys. J.* **212**, 60-70 (1977).
7. Higdon, J. C. *Astrophys. J.* **232**, 113-128 (1979).
8. Harding, A. K. *Astrophys. J.* (in the press).
9. Higdon, J. C. & Lingentel, R. E. *Astrophys. J. Lett.* **208**, L107-L110 (1976).
10. Strong, A. W., Wolfendale, A. W. & Dahanayake, C. *Mon. Not. R. astr. Soc.* **179**, 69-72 (1977).
11. Hartman, R. C. *et al.* in *The Structure and Content of the Galaxy and Galactic Gamma Rays* (eds Fichtel, C. E. & Stecker, F. W.) 15-25 (NASA CP-002, U.S. Government Printing Office, Washington DC, 1977).
12. Kniffen, D. A. *et al.* *Nature* **251**, 397-398 (1974).
13. Thompson, D. J., Fichtel, C. E., Kniffen, D. A. & Ögelman, H. *Astrophys. J. Lett.* **200**, L79-L82 (1975).
14. Bennett, K. *et al.* *Astr. Astrophys.* **61**, 279-284 (1977).
15. Lichti, G. G. *et al.* *Non-Solar Gamma Rays* (COSPAR Symp.) (eds Cowsik, R. & Wills, R. D.) 49-53 (Pergamon, New York, 1980).
16. Harding, A. K., Tademaru, E. & Esposito, L. W. *Astrophys. J.* **225**, 226-236 (1978).
17. Harding, A. K. *Astrophys. J.* (in the press).
18. Ögelman, H., Fichtel, C. E., Kniffen, D. A. & Thompson, D. J. *Astrophys. J.* **209**, 584-591 (1976).
19. Swenborg, B. N. *et al.* *Astrophys. J.* (in the press).
20. Weisberg, J. M., Rankin, J. & Boriakoff, V. *Astr. Astrophys.* **88**, 84-93 (1980).
21. Ables, J. G. & Manchester, R. N. *Astr. Astrophys.* **50**, 177-184 (1976).
22. Harding, A. K. *Pulsars* (eds Sieber, W. & Wielebinski, R.) (Reidel, Dordrecht, in the press).
23. Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215**, 885-896 (1977).
24. Stecker, F. W. *Astrophys. J.* **185**, 499-504 (1973).
25. Stecker, F. W. *Astrophys. Space Sci.* **6**, 377-389 (1970).
26. Stecker, F. W. *Astrophys. J.* **228**, 919-927 (1979).
27. Ganguli, S. N. & Sreekantan, B. V. *J. Phys. A*, **9**, 311-321 (1976).
28. Badhwar, G. D. & Stephens, S. A. *Proc. 15th Int. Cosmic Ray Conf. Plovdiv* **1**, 198-203 (1977).
29. Stecker, F. W. in *Origin of Cosmic Rays* (eds Osborne, J. L. & Wolfendale, A. W.) 267-334 (Reidel, Dordrecht, 1975).
30. Piccinotti, G. & Bignami, G. F. *Astr. Astrophys.* **52**, 69-75 (1976).
31. Daugherty, J. K., Hartman, R. C. & Schmidt, P. J. *Astrophys. J.* **198**, 493-505 (1975).
32. Cesarsky, D. J., Paul, J. A. & Shukla, P. G. *Astrophys. Space Sci.* **59**, 73-83 (1978).
33. Schlickeiser, R. & Thielheim, K. O. *Mon. Not. R. astr. Soc.* **182**, 103-109 (1978).
34. Strong, A. W., Wolfendale, A. W., Bennett, K. & Wills, R. D. *Mon. Not. R. astr. Soc.* **182**, 751-760 (1978).
35. Fichtel, C. E., Simpson, G. A. & Thompson, D. J. *Astrophys. J.* **222**, 833-849 (1978).
36. Lebrun, F. & Paul, J. A. *Proc. 16th Int. Conf. Cosmic Rays* 13-18 (1979).
37. Taylor, J. H. in *The Large-Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 119-123 (Reidel, Dordrecht, 1979).
38. Stecker, F. W. *Phys. Rev. Lett.* **35**, 188-191 (1975).
39. Stecker, F. W. *Nature* **260**, 412-414 (1976).
40. Stecker, F. W. & Jones, F. C. *Astrophys. J.* **217**, 843-858 (1977).
41. Stecker, F. W. *Am. Sci.* **66**, 570-576 (1978).
42. Stecker, F. W. *The Large Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 475-782 (Reidel, Dordrecht, 1979).

## Changes in orientation and appearance of the 100 m arc s radio structure in SS433

R. T. Schilizzi\*, G. K. Miley†, J. D. Romney‡ & R. E. Spencer§

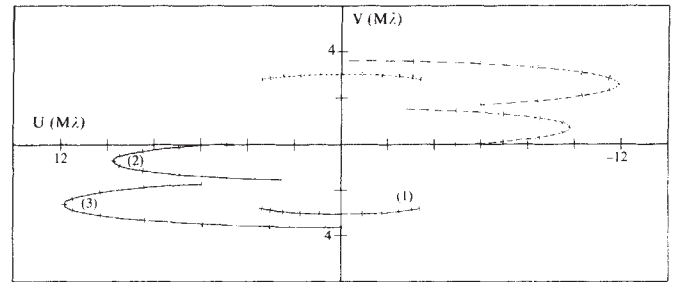
\* Netherlands Foundation for Radio Astronomy, Radio Sterrenwacht, Dwingeloo, The Netherlands

† Sterrewacht Leiden, Leiden, The Netherlands

‡ Max-Planck-Institut für Radioastronomie, Bonn, FRG

§ University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The discovery<sup>1</sup> of radio structure on a scale of 10–50 m arc s in SS433 has prompted us to monitor the object with the European VLBI network. The first observations in this programme were made for periods of 9 and 10 h at a wavelength of 6 cm on 9 January (JD2444248) and 1–2 June (JD2444392.5) 1980. Our results, together with the observations of Walker *et al.*<sup>2</sup> at an earlier epoch, show that the compact radio structure of SS433 changes orientation in a manner consistent with the 164-day precession period inferred from the optical emission lines, but with a phase delay of  $5 \pm 4$  days. The shape of the radio structure is elongated, non-linear and changes from one observing epoch to the other.



**Fig. 1** Loci of the east-west (U) and north-south (V) components of the projected baselines for our 6-cm observations of SS433. The baselines shown are those for which the source was detected: (1) Westerbork–Effelsberg; (2) Chilbolton–Effelsberg; (3) Knockin–Effelsberg.

Signals were recorded with the standard MkII VLBI system (bandwidth 2 MHz) and the cross correlations were performed at the Max-Planck-Institut für Radioastronomie in Bonn. Further observational details are given in Table 1. Useful data were obtained on the Westerbork–Effelsberg baseline in both sessions, on the Chilbolton–Effelsberg baseline for 2 h in January (the source was not detected during the rest of the observation) and on the Knockin (Jodrell Bank)–Effelsberg baseline for the full track in June. Loci of projected baselines for the various interferometers are shown in Fig. 1.

Measurements of calibrator sources were made at regular intervals throughout network periods encompassing the SS433 observations. Based on the flux densities measured at Westerbork and Effelsberg, these data were used to calibrate the performance of the individual interferometers<sup>3</sup>. In the June session, the unresolved source 2134+004 (ref. 4) at declination  $0.4^\circ$  was observed at several intervals during the first and last 3 h of its apparition, to calibrate directly the changing gain, aperture blockage and spillover effects of the various antennas as a function of elevation for SS433 (at dec  $4.9^\circ$ ). It was assumed that any differences in the general atmospheric conditions were negligible between observations of SS433 and 2134+004. The resulting calibration factors had an r.m.s. scatter of  $\sim 5\%$ . The point-to-point amplitude scatter for 2 min coherent integrations on the calibrator sources was  $\sim 3\%$ .

The flux densities of SS433 at 6 cm at the epochs of our measurements have been interpolated from measurements at 11 and 3.7 cm carried out by Johnston *et al.* (personal communication) assuming that the spectrum is a power law<sup>5</sup>. At these epochs, SS433 is classified as being in an inactive phase by Johnston *et al.*, meaning that the radio flux was changing by  $<20\%$  over a period of days. There is evidence that on 1–2 June, the flux density may have decreased by 8% during the period of our observations. We see little evidence of this in our visibility data, and it seems likely that variations of the visibility amplitude with hour angle are due to changes in the projected source structure rather than variations in the flux density.

The calibrated visibility amplitudes for the two sessions are shown in Fig. 2. Most of the information on source structure comes from the Westerbork–Effelsberg baseline, but the longer baselines provide important constraints on the structure. Although there are insufficient data for hybrid maps to be made, the relatively simple fringe amplitude curves allow several basic conclusions to be drawn about the structure and its variation. These conclusions are supported by the results of fitting simple models consisting of one or two barely resolved components and one or two extended elliptical gaussian components to the available data at each of the two epochs. Table 2 gives the parameters of the models and the range of values of each parameter which are consistent with the data. Curves representing the visibility amplitudes of these models are also shown in Fig. 2, and pictorial representations of the models from our data are given in Fig. 3.

The following deductions about the structure can be made.

(1) The size of the amplitude variations on these baselines



**Table 1** Some parameters of the observations

| Date           | UT        | Telescopes                             | Frequency (MHz) | Flux density (mJy) |
|----------------|-----------|----------------------------------------|-----------------|--------------------|
| 9 January 1980 | 0600–1500 | Westerbork<br>Effelsberg<br>Chilbolton | 4,997           | 500 ± 20           |
| 1–2 June 1980  | 2045–0700 | Westerbork<br>Effelsberg<br>Knockin    | 5,008.8         | 330 ± 30           |

The polarization on the sky was RCP in January and LCP in June.

shows the existence of structure on the 100–200 m arc s scale in both January and June.

(2) In January, the source was completely resolved on the longer projected spacings of the Chilbolton–Effelsberg baseline. The data shown for this baseline are for spacings similar to the Westerbork to Effelsberg baseline at extreme hour angles (see Fig. 1). In June, a barely resolved component of ~100 mJy is evident on the Knockin–Effelsberg baseline and on the Westerbork–Effelsberg baseline for the last 5 h.

(3) The projected baseline at which the peak fringe amplitude occurred for the Westerbork–Effelsberg interferometer differed in position angle (PA) by ~16° between the two observations. This indicates that the dominant PA of the 100–200 m arc s structure changed from  $98^\circ \pm 3^\circ$  on 9 January (when it was aligned with the bulges in W50) to  $114^\circ \pm 3^\circ$  on 1–2 June.

The following four points refer to the Westerbork–Effelsberg interferometer.

(4) An inflexion near IHA-6.5 h in the June curve, together with the best fitting model listed in Table 2, suggests that although the major part of the June emission was aligned in PA 114°, there is weaker emission still present in the ‘fundamental’ 98° PA.

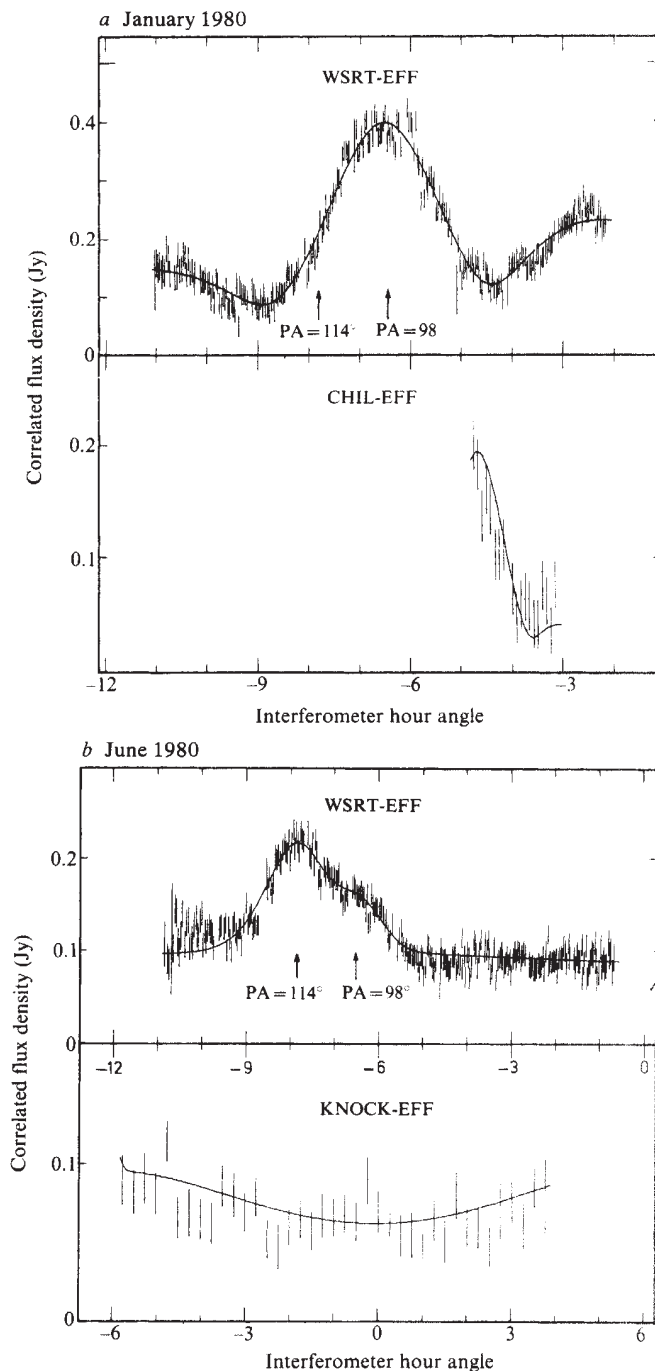
(5) The width of the peaks in both curves, as well as their asymmetry, indicate that the structure is non-linear and very elongated.

(6) The difference in shape between the two curves indicates that the general appearance of the radio source changed appreciably between the two observations.

(7) The amplitudes of the maxima indicate that at both epochs the flux density from the compact ( $\leq 150$  m arc s) structure was ~100 mJy less than the total flux density measured for SS433. This ‘missing’ 100 mJy can reasonably be ascribed to the larger (~3 arc s) structure known to exist from other measurements<sup>6,7</sup>.

The change in orientation of the compact radio structure may be compared with the predictions of models for SS433 derived from optical observations. On the basis of Abell and Margon’s<sup>8</sup> model, the PA of the precessing jets seen in projection on the sky is expected to vary approximately sinusoidally with time, with a peak-to-peak amplitude of  $41.4^\circ$  and with the same period as the optical emission lines if the jets describe a cone of half angle  $20.4^\circ$  about a mean inclination of  $79.7^\circ$  to the line of sight<sup>9,10</sup>. Alternatively, the PA has a saw-tooth variation with time (due to the  $180^\circ$  ambiguity in PA) with peak-to-peak amplitude of  $172.8^\circ$  if the angles are interchanged.

The position angles found from our data, and that of Walker *et al.*<sup>2</sup> on 12 March, 1979 ( $92^\circ$ , range  $85^\circ$ – $105^\circ$ ) were compared with the expected behaviour on Abell and Margon’s model and found consistent with precession with the smaller cone angle, as predicted by Begelman *et al.*<sup>11</sup>. If the axis of precession projects onto the sky in position angle  $98^\circ$  (corresponding to the bulges in W50<sup>10,12,13</sup>) then the PA of radio emission on the scale of 100 m arc s would lag the optical position angle by  $5 \pm 4$  days, and would be phased so that the PA was decreasing in June. Although probably fortuitous, this agrees with the expected delay of about 4.5 days for the radio emitting clouds to reach an angular distance of 50 m arc s, assuming a velocity of  $0.26c$  (ref. 4) and a distance<sup>10,12,14</sup> of SS433 of 4 kpc (50 m arc s corresponds to a length of  $6 \times 10^{15}$  cm at a distance of 4 kpc).



**Fig. 2** The visibility amplitudes for SS433 at 6 cm, as a function of interferometer hour angle. *a*, January 1980,  $S_T = 0.5$  Jy; *b*, June 1980,  $S_T = 0.33$  Jy. Arrows indicate the hour angles at which a maximum in visibility amplitude is expected for PA  $98^\circ$  and  $114^\circ$ . WSRT, Westerbork; EFF, Effelsberg; CHIL, Chilbolton; KNOCK, Knockin. Errors bars are  $2\sigma$ .

Our observations provide some evidence for extension of the radio source over a range in PA of  $\pm 5^\circ$  to  $\pm 10^\circ$  about the dominant PA. The presence of multiple-peaked spectral features have been noted<sup>9,15</sup> in optical data which could be evidence of remnant emission from previous positions of the beam during its precessional motion. It is not known whether the nonlinearity in the compact radio structure is also due to remnant emission and whether it corresponds to the possible optical remnant emission. Any correspondence will depend on the mechanism for ejecting material into the beams, and also on the lifetimes of the radiating electrons in the different wavelength domains. These factors will affect the persistence of the emission. In the radio domain, the non-thermal integrated spectrum and the high

Table 2 Parameters of the models

| Component             | Flux<br>(mJy) | Radial distance<br>(m arc s) | PA<br>(deg) | Angular size<br>(m arc s) | Component PA<br>(deg) |
|-----------------------|---------------|------------------------------|-------------|---------------------------|-----------------------|
| <i>a</i> January 1980 |               |                              |             |                           |                       |
| 1                     | 89 ± 25       | 0                            | 0           | (92 ± 12) × (1 ± 1)       | 101 ± 1               |
| 2                     | 219 ± 1       | 10 ± 1                       | 99 ± 3      | (22 ± 1) × (1 ± 1)        | 81 ± 2                |
| 3                     | 101 ± 3       | 65 ± 1                       | 99 ± 1      | (22 ± 1) × (1 ± 1)        | 89 ± 4                |
| <i>b</i> June 1980    |               |                              |             |                           |                       |
| 1                     | 134 ± 5       | 0                            | 0           | (163 ± 5) × (12 ± 2)      | 119 ± 1               |
| 2                     | 99 ± 1        | 0                            | 0           | (6.6 ± 0.5) × (5 ± 0.5)   | 98 ± 1                |
| 3                     | 56 ± 5        | 0                            | 0           | (164 ± 5) × (4 ± 1)       | 98 ± 1                |

The columns give the component number, the flux density, the radial distance from component 1, the position angle of the location of the component (N to E) with respect to component 1, the major and minor axis, and the position angle of the component. The errors given are formal errors from the model-fitting process; the true errors are probably larger.

brightness temperatures of the compact components, ranging up to  $10^9$  K (ref. 2), imply that the emission is synchrotron radiation from relativistic electrons. The optical emission line system from the beams seems to arise from gas at a temperature of  $<10^4$  K (ref. 11). There are arguments<sup>15</sup> that the lifetime of the optical line emission is  $\leq 10$  days, but in the radio case there is not enough structural information on which to base calculations of synchrotron lifetimes. These relative lifetimes will also have some bearing on the lag of the rotation of the main axis of the radio emission behind that of the optical.

Although the optical data predict that during the January and June observations, the beams of SS433 lie within  $10^\circ$  of the plane of the sky, the detailed radio structures are different. Further data are required before patterns (if any exist) can be found in the variation of the structure. Radio flares will affect the structure for some time afterwards; note that the January data were taken 8 days after flaring activity in SS433 has ceased while the June data were taken 11 days before the end of the same long

period of quiescent behaviour (K. J. Johnston, personal communication). Perhaps the enhanced compact component in the June data can be identified with a pre-flare buildup of relativistic particles.

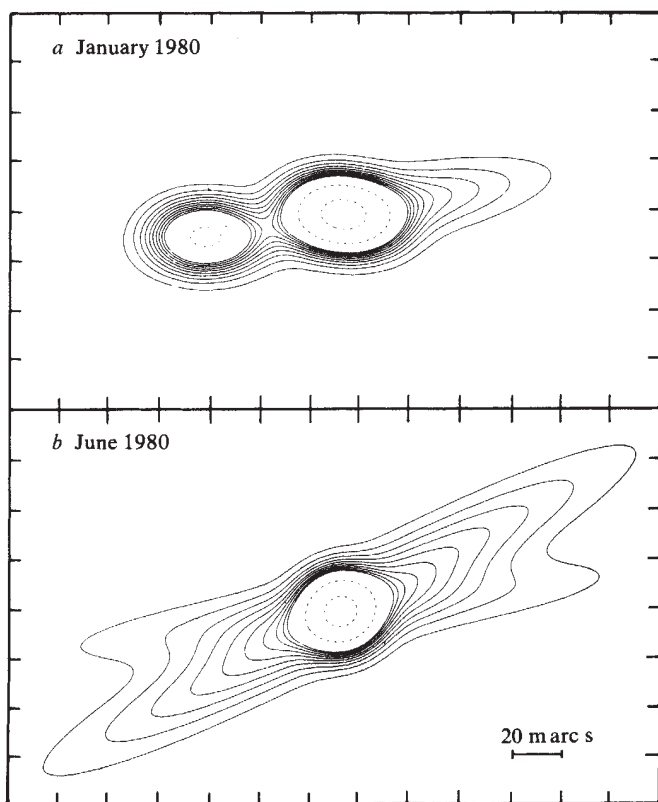
We have shown that the 100–200 m arc s structure in SS433 varies in orientation. Although the behaviour is consistent with the precessing beams model, further observations are needed to establish that the change in position angle is indeed periodic. Future comparison of variations in the radio structure and in the optical line emission may well provide unique information about the SS433 phenomenon.

We thank the staffs of the various observatories for their assistance, in particular R. S. Booth and P. N. Wilkinson at the UK telescopes; R. W. Porcas, H. Blaschke and U. Stursberg for advice and assistance during the processing of the data; and K. Johnston for flux density information. The Westerbork Synthesis Radio Telescope is operated by the Netherlands Foundation for Radio Astronomy with the financial support of the Netherlands Organization for Pure Research (ZWO).

**Note added in proof:** Hjellming and Johnston<sup>16</sup> report extensive VLA observations of SS433 at seven epochs from 8 September 1979 to 20 June 1980. Their maps show an unresolved ( $\leq 100$  m arc s) core source plus elongated 'jet' components up to several arc s in extent. Several of their results bear on the present paper. (1) The PA of the arc s scale structure remains within  $\pm 20^\circ$  of  $100^\circ$ , which we have also deduced from our data with reference to the Abell and Margon model. (2) Proper motions observed for several features in the VLA maps have been extrapolated back to show that the radio jets are probably associated with the optical jets in the inner regions. There is some evidence from their data that the radio precession ephemeris lags that of the optical by 10–20 days with an error of the same magnitude. This is consistent with our result. (3) Assuming the core source is spherical, equipartition arguments lead Hjellming and Johnston to synchrotron loss lifetimes of  $10^9$  s, from which it is concluded that these losses cannot be significant in the core. The structure we derive for the central source is highly elongated implying higher magnetic fields and shorter synchrotron lifetimes. We do not believe that the structural information on the small scales is yet sufficiently good to confidently rule out one or another loss mechanism.

Received 15 December 1980; accepted 29 January 1981.

- Schilizzi, R. T., Norman, C. A., van Breugel, W. & Hummel, E. *Astr. Astrophys.* **79**, L26–L27 (1979).
- Walker, R. C. *et al. Astrophys. J.* (in the press).
- Cohen, M. H. *et al. Astrophys. J.* **201**, 249–255 (1975).
- Pauliny-Toth, I. I. K. *et al. Astr. J.* (submitted).
- Seaquist, E. R., Gilmore, W., Nelson, G. J., Payten, W. J. & Slee, O. B. *Astrophys. J. Lett.* **241**, L77–L81 (1980).
- Spencer, R. E. *Nature* **282**, 483–484 (1979).
- Gilmore, W. & Seaquist, E. R. *Astr. J.* **85**, 1486–1495 (1980).
- Abell, G. O. & Margon, B. *Nature* **279**, 701–703 (1979).
- Margon, B., Grandi, S. A. & Downes, R. A. *Astrophys. J.* **241**, 306–315 (1980).
- Geldzahler, B. J. thesis, Univ. Pennsylvania (1980).
- Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* **238**, 722–730 (1980).
- Geldzahler, B. J., Pauls, T. & Salter, C. J. *Astr. Astrophys.* **84**, 237–244 (1980).
- Seward, F., Grindlay, J., Seaquist, E. & Gilmore, W. *Nature* **287**, 806–808 (1980).
- Van Gorkom, J., Goss, W. M. & Shaver, P. A. *Astr. Astrophys.* **82**, L1–L2 (1980).
- Murdin, P., Clark, D. H. & Martin, P. G. *Mon. Not. R. astr. Soc.* **193**, 135–151 (1980).
- Hjellming, R. M. & Johnston, K. J. *Nature* **290**, 100–107 (1981).



**Fig. 3** Models for SS433 at 6 cm. *a*, 9 January 1980; *b*, 1–2 June 1980. There is a  $180^\circ$  ambiguity in the orientation of these diagrams because they were derived using amplitude data alone. North or south is up. The data have been smoothed with a 20 m arc s circular gaussian beam. The contour levels are: January 1.5, 3, 4, 5, ... 15% (solid), 45, 90% (dashed); June, 2, 4, 6, ... 20% (solid), 40, 80% (dashed).



## Acetonitrile in the stratosphere—implications from laboratory studies

H. Böhringer & F. Arnold

Max-Planck-Institut für Kernphysik, PO Box 103980, D6900 Heidelberg, FRG

The presence of acetonitrile ( $\text{CH}_3\text{CN}$ ) in the stratosphere has been tentatively inferred<sup>1</sup> from measurements of stratospheric positive ions around 36 km. As well as the expected proton hydrates  $\text{H}^+(\text{H}_2\text{O})_n$ , another ion family was observed to which was assigned the general form,  $\text{H}^+\text{X}_l(\text{H}_2\text{O})_m$ , where X is a molecule of mass  $41 \pm 1$  AMU. If these ions are formed from proton hydrates by reactions involving the unknown trace gas X, the latter must have a proton affinity of  $\geq 175$  kcal mol<sup>-1</sup> and a number density of the order of  $10^5$  cm<sup>-3</sup> around an altitude of 36 km. Two previously proposed identifications for X—NaOH and MgOH—have recently been excluded by new stratospheric ion composition data<sup>3,4</sup> which unambiguously determined the mass of X to be 41 AMU and failed to show  $\text{H}^+\text{X}_l(\text{H}_2\text{O})_m$  ions containing the Mg isotopes 25 and 26. Another suggestion<sup>1</sup> that X might be acetonitrile ( $\text{CH}_3\text{CN}$ ) has been criticized on the grounds that it is difficult to explain its presence in the stratosphere<sup>2</sup>. We now report laboratory studies of clustering equilibria for  $\text{H}^+(\text{CH}_3\text{CN})_l(\text{H}_2\text{O})_m$  ions, which further support identification of X as acetonitrile.

The measurements were made using a newly designed temperature-variable drift tube mass spectrometer which will be described in detail elsewhere<sup>5</sup>. In this apparatus the primary ions are produced in a separate ion source and are mass selected before they are introduced into the drift tube. The ions traverse the tube drifting in an electric field while they are subjected to reactions with reactant gases. Having passed the drift tube exit aperture the product ions are mass analysed. During the measurements described here the drift tube was operated at low electric fields ( $\leq 2$  V cm<sup>-1</sup>) so that the additional 'drift energy' of the ions was  $< 2\%$  of the thermal energy and only minor temperature corrections have been applied for this effect. The gas pressures applied to the drift tube ranged from 0.2 to 3 torr and the temperatures from 220 to 400 K. Hydrogen was used as buffer gas to which the reactant gas  $\text{CH}_3\text{CN}$  or  $\text{H}_2\text{O}$  was added in amounts of 0.2–10%. The second of the two gases was only added in traces if mixed ligand clusters were to be observed. Primary  $\text{H}_2^+$  ions were injected into the drift tube and rapidly converted to  $\text{H}_3^+$  by reactions with hydrogen. Proton transfer reactions of these ions led to the core ions  $\text{H}_3\text{O}^+$  and  $\text{CH}_4\text{CN}^+$ , which were subjected to further clustering and ligand exchange reactions. We assumed that chemical equilibrium had been established when a further increase of the buffer gas pressure and the reaction time did not change the ratio of product to precursor ion abundance. Absorption and desorption of the reactant gas at the walls of the drift tube made control of the reactant gas pressure difficult and introduced an estimated error of up to a factor of 3 in the measured equilibrium constant.

The  $\text{H}^+(\text{CH}_3\text{CN})_l(\text{H}_2\text{O})_m$  ions produced in the laboratory showed all the ion masses observed in the stratosphere by Arnold *et al.*<sup>1,3</sup>, and essentially all ion masses detected by Arijis *et al.*<sup>4,6</sup> at an altitude around 34 km (see Table 1). The measured equilibrium constants for association of  $\text{H}_2\text{O}$  and  $\text{CH}_3\text{CN}$  to different cluster ions are given in Fig. 1 for a temperature of 227 K (the ambient temperature during the stratospheric measurement in ref. 3). A comparison of the equilibrium constants for the proton hydrates with previous high-pressure ion source measurements<sup>7</sup> also given in Fig. 1 shows a fairly good agreement.

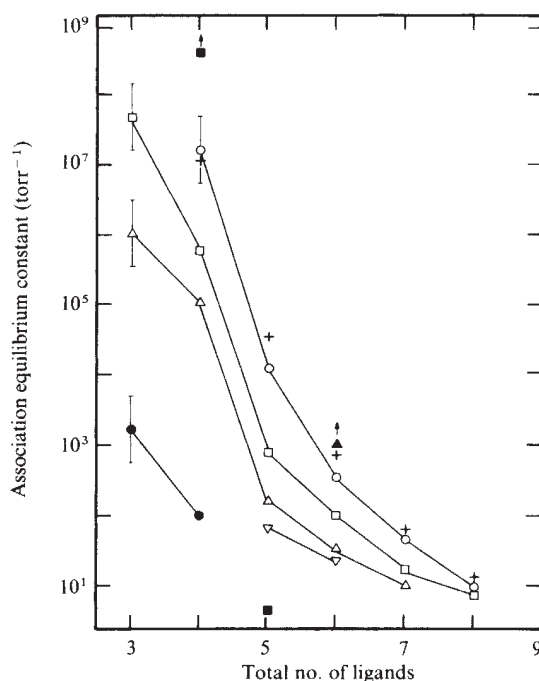
The results show that hydration equilibrium constants for heteromolecular clusters containing  $\text{CH}_3\text{CN}$  molecules are generally lower than those for  $\text{H}^+(\text{H}_2\text{O})_n$  ions and they decrease with increasing number of  $\text{CH}_3\text{CN}$  ligands contained in the

complex. The equilibrium constants for the  $\text{H}^+(\text{CH}_3\text{CN})_p$  ions with  $p \geq 3$  are remarkably low despite the large dipole moment of  $\text{CH}_3\text{CN}$ . Thus, it seems that the ligand interaction between  $\text{H}_2\text{O}$  and  $\text{CH}_3\text{CN}$  and among  $\text{CH}_3\text{CN}$  molecules is weaker than that between  $\text{H}_2\text{O}$  molecules which is dominated by hydrogen bonding. This is consistent with the lower heat of evaporation of  $\text{CH}_3\text{CN}$  compared with  $\text{H}_2\text{O}$ .

A striking exception, however, is found for the ion  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  which is the most stable ion containing four ligands and has a much larger equilibrium constant than the two precursors,  $\text{H}^+(\text{CH}_3\text{CN})_3$  and  $\text{H}^+(\text{CH}_3\text{CN})_2\text{H}_2\text{O}$ . Using collision-induced dissociation, we found that decomposition of  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  preferentially leads to the fragment  $\text{H}^+(\text{CH}_3\text{CN})_2\text{H}_2\text{O}$ , implying that the water molecule is bound more strongly than the acetonitrile molecules. Therefore, the high stability of this ion may be explained by a structural rearrangement, in which  $\text{H}_3\text{O}^+$  forms the core ion and the acetonitrile molecules are symmetrically attached to the three hydrogen atoms sharing the positive charge. A similarly enhanced stability was found for the ion  $\text{H}^+(\text{CH}_3\text{CN})_4(\text{H}_2\text{O})_2$  among the ions with six ligands. An analogous symmetric structure can be attributed to this ion in which the two water molecules attach to the proton leaving four, partly charged H-atom bonding sites for the acetonitrile molecules.

In the stratosphere, hydration is sufficiently fast to attain hydration equilibria for all positive ions. Therefore the hydration equilibria of stratospheric X-containing and laboratory  $\text{CH}_3\text{CN}$ -containing ions can be compared—as shown in Table 2—and a fairly good agreement is found. The laboratory results indicate that among the  $\text{H}^+(\text{CH}_3\text{CN})_3(\text{H}_2\text{O})_k$  family,  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  should be the only detectable atmospheric ion, which is, in fact, shown by the stratospheric observations.

The stratospheric observations of  $\text{H}_2\text{O}$ -clustering equilibria are therefore consistent with X being acetonitrile. The only



**Fig. 1** Measured equilibrium constants for  $\text{H}_2\text{O}$  and  $\text{CH}_3\text{CN}$  association reactions at 227 K. The equilibrium constants  $\geq 10^4$  (torr<sup>-1</sup>) had to be extrapolated from higher temperatures. For  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  and  $\text{H}^+(\text{CH}_3\text{CN})_4(\text{H}_2\text{O})_2$  only lower limits can be given. The results for the proton hydrates are compared with measurements of Kebarle *et al.*<sup>6</sup>. The estimated error of up to a factor of 3 valid for all measured values is indicated by error bars for the first points of the curves.  $\circ$ ,  $\text{H}^+(\text{H}_2\text{O})_n$ ;  $\square$ ,  $\text{H}^+\text{CH}_3\text{CN}(\text{H}_2\text{O})_m$ ;  $\triangle$ ,  $\text{H}^+(\text{CH}_3\text{CN})_2(\text{H}_2\text{O})_l$ ;  $\nabla$ ,  $\text{H}^+(\text{CH}_3\text{CN})_3(\text{H}_2\text{O})_k$ ;  $\bullet$ ,  $\text{H}^+(\text{CH}_3\text{CN})_p$ ;  $\blacksquare$ ,  $\text{H}^+\text{H}_2\text{O}(\text{CH}_3\text{CN})_q$ ;  $\blacktriangle$ ,  $\text{H}^+(\text{H}_2\text{O})_2-(\text{CH}_3\text{CN})_r$ ;  $+$ ,  $\text{H}^+(\text{H}_2\text{O})_n$  (ref. 6).

exception is the observed ion with mass  $136 \pm 1$  AMU. Assuming that the previous identification,  $\text{H}^+(\text{CH}_3\text{CN})_2(\text{H}_2\text{O})_3$ , is correct, the ion should be one or two orders of magnitude less abundant than observed. Therefore, this ion could result from reactions involving another trace gas, Y, and thus have the form  $\text{H}^+\text{Y}(\text{H}_2\text{O})_i$  or  $\text{H}^+\text{YX}(\text{H}_2\text{O})_j$ . If this cluster ion contains three or four ligands like the other observed heteromolecular clusters, the mass of the molecule Y should be  $99 \pm 1$ ,  $81 \pm 1$ ,  $76 \pm 1$  or  $58 \pm 1$  AMU. Possible candidates for Y are  $\text{CH}_3\text{NO}_3$  (mass 77),  $\text{C}_2\text{H}_5\text{CHO}$  (mass 58), and  $(\text{CH}_3)_2\text{CO}$  (mass 58) which all have proton affinities in excess of  $175 \text{ kcal mol}^{-1}$  and therefore should replace water molecules in proton hydrates in stratospheric conditions. Note that a recently detected<sup>4</sup> ion of minor abundance with mass 117 AMU may belong to the same hydrate ion family as the ion  $136 \pm 1$  AMU. If so, the mass of Y should be 99, 80, 75 or 57 which would exclude all identifications suggested above.

**Table 1** Observed stratospheric positive ions around 36 km altitude

| Mass        | Identification                               | Ion mass for<br>$\text{X} = \text{CH}_3\text{CN}$ | Measured abundance<br>(%) |        |
|-------------|----------------------------------------------|---------------------------------------------------|---------------------------|--------|
|             |                                              |                                                   | Ref. 3                    | Ref. 4 |
| 55          | $\text{H}^+(\text{H}_2\text{O})_3$           | 55                                                | 3.4                       | 5.5    |
| 73          | $\text{H}^+(\text{H}_2\text{O})_4$           | 73                                                | 26.8(21.3)*               | 29.6   |
| 78          | $\text{H}^+\text{X}(\text{H}_2\text{O})_2$   | 78                                                | 7.2                       | 8.4    |
| 91          | $\text{H}^+(\text{H}_2\text{O})_5$           | 91                                                | 8.9(17.8)*                | 2.8    |
| 96          | $\text{H}^+\text{X}(\text{H}_2\text{O})_3$   | 96                                                | 19.6                      | 44.4   |
| 101         | $\text{H}^+\text{X}_2(\text{H}_2\text{O})_2$ | 101                                               | 10.7                      | 3.5    |
| 114         | $\text{H}^+\text{X}(\text{H}_2\text{O})_4$   | 114                                               | —                         | 0.9    |
| $119 \pm 1$ | $\text{H}^+\text{X}_2(\text{H}_2\text{O})_2$ | 119                                               | 16.1                      | 4.9    |
| $136 \pm 1$ | $\text{H}^+\text{X}_2(\text{H}_2\text{O})_3$ | 137                                               | 4.4                       | —      |
| $141 \pm 1$ | $\text{H}^+\text{X}_3\text{H}_2\text{O}$     | 142                                               | 2.9                       | —      |

\* Values in parentheses are corrected for dissociation (see ref. 3).

Finally, implications of the measured  $\text{CH}_3\text{CN}$  clustering equilibria will be considered. Arnold *et al.*<sup>3</sup> have applied a steady-state treatment of the reactions converting the proton hydrates to cluster ions, containing 1, 2 and 3 X molecules, to their stratospheric ion composition data. They found that the ion family containing three X ligands is one order of magnitude less abundant than expected, assuming equal reaction rates for all three conversion steps. This was tentatively explained by a process converting the only observed ion with three  $\text{CH}_3\text{CN}$  molecules,  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ , back to cluster ions containing two  $\text{CH}_3\text{CN}$  ligands. The process should be about 10 times faster than ion-ion recombination in the stratosphere around 36 km. The laboratory studies revealed that the  $\text{CH}_3\text{CN}$  molecules are less strongly bound than  $\text{H}_2\text{O}$  in the  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  complex. Therefore, two processes could be responsible for the backward conversion: thermal decomposition or ligand exchange of  $\text{CH}_3\text{CN}$  for a  $\text{H}_2\text{O}$  molecule. For the thermal dissociation, a free enthalpy value of  $\Delta G \approx 13.8 \text{ kcal mol}^{-1}$  is calculated for 227 K and a gas density of  $4 \times 10^{16} \text{ cm}^{-3}$  to account for the observed ion abundances, while for the ligand exchange reaction  $\Delta G \approx 5.8 \text{ kcal mol}^{-1}$  is obtained for the concentration ratio  $[\text{H}_2\text{O}]/[\text{X}] = 3 \times 10^6$  (ref. 3). The corresponding values derived from the laboratory measurements for the  $\text{CH}_3\text{CN}$ -containing clusters are  $\Delta G \geq 12 \text{ kcal mol}^{-1}$  and  $\Delta G \geq 3.6 \text{ kcal mol}^{-1}$ . A thermal dissociation or ligand exchange reaction affecting the  $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$  ion is also observed in the laboratory studies at large  $[\text{H}_2\text{O}]/[\text{CH}_3\text{CN}]$  ratios ( $\gg 10^3$ ). We therefore conclude that the  $\Delta G$  values for both processes hardly exceed the lower limits given above by a large factor, which agrees well with the findings from the atmospheric observations.

Thus the observed masses and abundances of positive stratospheric ions can be well explained by proton hydrates and  $\text{H}^+(\text{CH}_3\text{CN})_i(\text{H}_2\text{O})_m$  ion clusters. The ions  $136 \pm 1$  and 117 AMU which do not fit into this scheme might reflect the influence of a stratospheric trace gas other than acetonitrile.

Until now, no source of stratospheric acetonitrile has been known. The sharp decrease in fractional abundances of X-

**Table 2** Relative comparison of hydration equilibria of laboratory and stratospheric ions

| Abundance ratio     | Laboratory data                            | Ref. 3 | Ref. 4 |
|---------------------|--------------------------------------------|--------|--------|
| [96]/[78]           | $1.7 \times 10^{-11} [\text{H}_2\text{O}]$ | 2.7    | 5.3    |
| [119]/[101]         | $3.5 \times 10^{-12} [\text{H}_2\text{O}]$ | 1.5    | 1.4    |
| [91]/[73]           | $3.5 \times 10^{-13} [\text{H}_2\text{O}]$ | 0.84   | 0.1    |
| [114]/[96]          | $2.2 \times 10^{-14} [\text{H}_2\text{O}]$ | —      | 0.02   |
| [136 $\pm$ 1]/[119] | $4.7 \times 10^{-15} [\text{H}_2\text{O}]$ | 0.27   | —      |

Ions are referred to by their mass number.

containing ions above 40 km observed in rocket experiments<sup>8</sup> indicates that the source of X should be at heights below about 35 km. Further ion composition measurements covering a larger height range may help locate this source and provide additional thermodynamic data on X-containing ions which can be compared with our laboratory results.

We thank Professor E. E. Ferguson for helpful discussions. Part of the project was funded by the Bundesminister für Forschung und Technologie through GSF.

*Note added in proof:* The kinetics of the switching reactions of proton hydrates with  $\text{CH}_3\text{CN}$  has been studied by Smith *et al.*<sup>9</sup>

Received 1 December 1980; accepted 12 February 1981.

1. Arnold, F., Böhringer, H. & Henschen, G. *Geophys. Res. Lett.* **5**, 653–656 (1978).
2. Ferguson, E. E. *Geophys. Res. Lett.* **5**, 1035–1038 (1978).
3. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* **29**, 185–193 (1981).
4. Aijis, E., Nevejans, D. & Ingels, I. *Nature* **288**, 684–686 (1980).
5. Böhringer, H. & Arnold, F. *Int. J. Mass Spectrom. Ion Phys.* (submitted).
6. Aijis, E., Ingels, J. & Nevejans, D. *Nature* **271**, 642–644 (1978).
7. Kebarle, P., Searles, S. K., Zolla, A., Scarborough, J. & Arshadi, M. *J. Am. chem. Soc.* **89**, 6393–6399 (1967).
8. Arnold, F., Krankowsky, D. & Marien, K. H. *Nature* **267**, 30–32 (1977).
9. Smith, D., Adams, N. G. & Aige, E. *Planet. Space Sci.* (in the press).

## Unusual *PT* dependence of thermal conductivity for a clathrate hydrate

R. G. Ross, P. Andersson & G. Bäckström

Department of Physics, Umeå University, S-901 87 Umeå, Sweden

**Molecular clathrate solids<sup>1</sup> are crystals with peculiar and interesting properties. Their characteristic feature<sup>2</sup> is that one chemical species, the 'host', forms a strongly bonded structure, which encages individual molecules of a second 'guest' species. There is virtually no chemical bonding between host and guest molecules. The guests are, nevertheless, firmly constrained by high energy barriers. X-ray and neutron diffraction experiments have shown that clathrates have a periodic structure typical of crystalline solids. A clathrate hydrate<sup>2</sup> is a host substance consisting of  $\text{H}_2\text{O}$  molecules which form an ice-like, hydrogen-bonded structure. Huge geological deposits of clathrate hydrates containing natural gas are known to exist<sup>2,3</sup>—these deposits may be exploitable, and in searching for clathrate hydrates knowledge of the thermal conductivity,  $\lambda$ , will be important. We present here a preliminary account of the first measurements of  $\lambda$  versus pressure (*P*) and temperature (*T*) for a clathrate hydrate. We studied a clathrate hydrate encaging molecules of tetrahydrofuran. Thermal conductivity was measured using the transient hot-wire method<sup>4</sup>. The accuracy of the measurements was 3% and the precision 1%. The temperature dependence was found to be unusual in that  $\lambda(T)$  has a positive slope (Fig. 1), a feature never before observed with crystalline organic materials. Furthermore, when  $\lambda$  was measured versus *P*, it was found unexpectedly that  $\lambda(P)$  was constant—behaviour never before reported for a non-metallic material.**

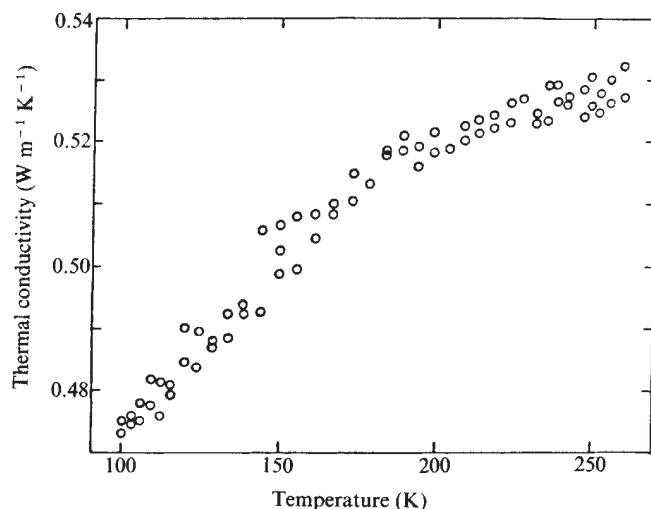
The temperature dependence almost universally observed at  $T > 100 \text{ K}$  with crystalline, non-metallic solids is  $\lambda(T) \propto T^{-1}$ . This negative slope agrees with the theory of phonon-phonon scattering by the simplest umklapp mechanism. Positive slopes



of  $\lambda(T)$  are encountered for non-crystalline materials in the glass state. Here topological structural disorder is considered to provide the dominant phonon scattering mechanism. Disorder fixes the phonon mean free path, and the positive slope of  $\lambda(T)$  is attributed to variation of the heat capacity in the well-known formula for  $\lambda$  due to Debye. A few inorganic crystalline solids have positive slopes of  $\lambda(T)$ , that is  $\text{YB}_{68}$ ,  $\text{KFeF}_3$  and  $\text{FeAl}_2\text{O}_4$ , and structural disorder has been invoked to account for this behaviour<sup>5-7</sup>. Orientationally disordered plastic crystal phases<sup>8</sup> are intermediate cases, showing a negative slope of  $\lambda(T)$  but with much reduced steepness:  $\lambda \propto T^{-0.2}$ .

All non-metallic materials previously investigated exhibit a finite slope of  $\lambda(P)$ , usually positive, as would be expected. A few cases of negative slope have, however, recently been discovered, one being ordinary ice<sup>9</sup>. The magnitudes of the slopes are most conveniently compared in terms of  $g = \partial \ln \lambda / \partial \ln \rho$ , where  $\rho$  is the mass density. For most cases  $g$  ranges from 3 to 10, whereas for exceptional materials  $g = -2.6$  (ice Ih)<sup>9</sup> and  $g = -6.2$  ( $\text{NH}_4\text{F}$ )<sup>10,11</sup>. From the compressibility of a propane clathrate hydrate<sup>2,3</sup>, on the other hand, we estimate  $|g| < 0.2$ , which is an order of magnitude smaller than for other substances.

The thermal conductivity of the clathrate was found to be only 15% greater than for the corresponding liquid solution. This is in qualitative agreement with the only previous measurement for a propane hydrate<sup>3</sup>. The increase in  $\lambda$  when going from liquid to an orientationally disordered plastic crystal phase<sup>8</sup> is typically 5%, whereas a dramatic increase is usually recorded for a normal crystal<sup>12</sup>.



**Fig. 1** Isobaric variation with temperature of thermal conductivity,  $\lambda$ , of (tetrahydrofuran)·17.05( $\text{H}_2\text{O}$ ) clathrate hydrate at a pressure of 0.1 GPa.

Although the tetrahydrofuran clathrate hydrate is crystalline, there is in fact some structural disorder<sup>2</sup>. Recent neutron inelastic scattering results<sup>13</sup> indicate reorientational motion of the guest molecules with an activation energy of  $\sim 0.01$  eV. At temperatures of 100–300 K there would thus be frequent reorientation of the guest molecules, which might provide a phonon scattering mechanism. We propose that this dynamical structural disorder causes a dominant thermal resistivity in our clathrate hydrate. From the pressure variation of  $\lambda$  we conclude that the total resistive mechanism is virtually insensitive to changes in density, but we cannot yet exclude that this constancy is a combined effect of positive and negative contributions from the main scattering mechanisms.

The phase diagram for mixtures of tetrahydrofuran and  $\text{H}_2\text{O}$  versus  $P$ ,  $T$  and composition is well known<sup>14</sup>. The so-called structure II clathrate hydrate exists in a very narrow region of composition. At atmospheric pressure it occurs at a composition of (tetrahydrofuran)·17( $\text{H}_2\text{O}$ ) with congruent melting. We prepared our specimens by pouring a liquid solution of double-

distilled  $\text{H}_2\text{O}$  and spectroscopic-grade tetrahydrofuran into the Teflon measurement cell of a piston-cylinder high pressure apparatus. On decreasing the temperature at 0.01 GPa freezing occurred after 5 K of supercooling and was detected by a marked thermal arrest at 277.0 K in agreement with the phase diagram<sup>14</sup>. Such thermal behaviour during solidification excludes the possibility of glass formation. The observed discontinuity of  $\lambda$  at the arrest temperature corroborates the conclusion that a first-order transition was induced on cooling. Experiments were carried out at four different compositions from 16.59 to 17.55 moles of  $\text{H}_2\text{O}$ , and the variation of  $\lambda$  versus  $T$  and  $P$  remained the same, although the absolute values of  $\lambda$  differed by as much as 20% between the extreme cases. It is thus unlikely that the variation of  $\lambda$  reported here could have been caused by phase separation in response to changes in  $P$  and  $T$ .

It is reasonable to assume that natural-gas clathrate hydrates would behave similarly to the substance investigated here. One could thus expect  $\lambda$  to be small and to change little over geologically relevant ranges of temperature and pressure.

We thank Dr D. W. Davidson for suggesting this investigation and Dr G. A. Slack for useful discussion. The work was financially supported by the Swedish Natural Science Council.

Received 8 December 1980; accepted 27 January 1981.

1. Parsonage, N. G., & Staveley, L. A. K. *Disorder in Crystals* (Clarendon, Oxford, 1978).
2. Davidson, D. W. in *Water: A Comprehensive Treatise* Vol. 2 (ed. Franks, F.) 115–234 (Plenum, New York, 1973).
3. Stoll, R. D. & Bryan, G. M. J. *geophys. Res.* **84**, 1629–1634 (1979).
4. Ross, R. G., Andersson, P. & Bäckström, G. *Molec. Phys.* **38**, 377–385 (1979).
5. Slack, G. A., Oliver, D. W. & Horn, F. H. *Phys. Rev.* **B4**, 1714–1720 (1971).
6. Suemune, Y. *J. phys. Soc. Jap.* **19**, 2234 (1964).
7. Slack, G. A. *Phys. Rev.* **134**, A1268–A1280 (1964).
8. Andersson, P. & Ross, R. G. *Molec. Phys.* **39**, 1359–1368 (1980).
9. Andersson, P., Ross, R. G. & Bäckström, G. *J. phys. C: Solid State Phys.* **13**, L73–L76 (1980).
10. Ross, R. G. & Sandberg, O. *J. phys. C: Solid State Phys.* **11**, 667–672 (1978).
11. Swenson, C. A. & Tedeschi, J. R. *J. chem. phys.* **40**, 1141–1143 (1964).
12. Slack, G. A. *Solid State Phys.* **34**, 1–71 (1979).
13. Wegener, W., Vanderhaeghen, J., Hautecler, S., Legrand, E. & van Gervan, L. in *Neutron Inelastic Scattering 1977*, Vol. 1, 415–433 (IAEA, Vienna, 1978).
14. Dyadin, Yu. A., Kuznetsov, P. N., Yakovlev, I. I. & Pyrinova, A. V. *Doklady. Chem.* **208**, 9–12 (1973).

## Destruction of polychlorodibenzo-*p*-dioxins

D. C. Ayres

Department of Chemistry, Westfield College, Hampstead, London NW3 7ST, UK

The release of 2,4,5-trichlorophenol (2,4,5-T) containing 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) at Seveso<sup>1</sup> highlighted the need for an efficient control procedure for the highly toxic<sup>2</sup> dioxin. The concomitant formation of the dioxin and the trichlorophenol during hydrolysis of *sym*-tetrachlorobenzene has led to demands for restrictions on the use of 2,4,5-T and related herbicides<sup>3</sup>. An environmental hazard is also presented by pentachlorophenol<sup>4</sup> which includes the more highly chlorinated dioxins, in particular octachlorodibenzo-*p*-dioxin (OCDD). The use of polychlorophenol formulations has been banned in Sweden on the grounds<sup>5</sup> that heat-stable chlorodioxins are formed during incineration; this also points to the need for degradation by chemical means. Dehalogenation of polychlorodibenzodioxins (PCDD) can be accomplished by photolysis in laboratory conditions<sup>6</sup>, but material adsorbed on soil is little affected<sup>7</sup>. Cleavage of the ether linkages with the formation of halophenols may be achieved by treatment with strong acids or quaternary ammonium salts<sup>8</sup>. The dibenzodioxin nucleus, is however, rather resistant to chemical attack<sup>9</sup> and in the absence of a suitable control procedure the use of 2,4,5-T has been restricted in the US by the Environmental Protection Agency<sup>10,11</sup>. I now report the oxidative degradation of TCDD and related compounds by ruthenium tetroxide.

Ruthenium tetroxide is a powerful oxidizing agent which reacts with a wide range of organic substances<sup>12</sup> including those aromatic compounds which are oxidized slowly, or not at all, by permanganate<sup>13</sup>. It may be safely used in solution in water or in organic solvents with no nucleophilic character, such as chloroform, nitromethane or carbon tetrachloride.

In a preliminary experiment we showed that 2,7-dichlorobenzo-*p*-dioxin (DCDD) is destroyed by ruthenium tetroxide in carbon tetrachloride solution on standing overnight at ambient temperature. The progress of this and related oxidations may be monitored by thin layer or preferably by gas chromatography, where a 2-m column of 5% OV1 on Chromosorb 750 was operated at 195 °C, with a nitrogen flow of 50 ml min<sup>-1</sup> giving a retention time of 162 s.

Subsequent work on TCDD was carried out in collaboration with J. Harrison at the Chemical Defence Establishment, Porton Down. The progress of the oxidation of TCDD ( $C = 72 \mu\text{g ml}^{-1}$ ) at ambient temperature in carbon tetrachloride containing the tetroxide ( $2 \times 10^3 \mu\text{g ml}^{-1}$ , about 10 equiv.) was monitored at intervals by stopping aliquots with benzene (2 ml). Residual concentrations were determined by comparison of their electron response with that of a standard solution of TCDD ( $C = 60 \mu\text{g ml}^{-1}$ ) with the GLC operating at 230 °C. The half life of 560 min was evaluated from the linear pseudo-first order plot of

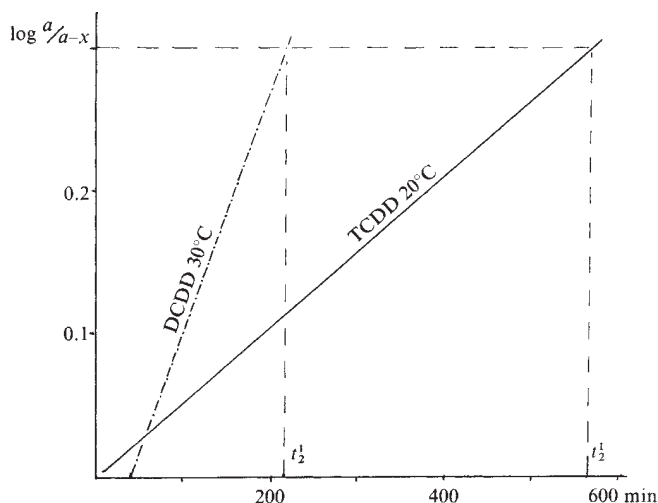


Fig. 1 Rates of oxidation of polychlorodibenzodioxins by ruthenium tetroxide.

log (proportion residual)<sup>-1</sup> to time (Fig. 1). The rate of oxidation was considerably increased at higher temperatures: in carbon tetrachloride at 70 °C and with the same concentration of reagents the half life of TCDD was <15 min.

For general guidance, within this group of compounds, the temperature coefficient for the oxidation of 2,7-DCDD was determined. The rate of oxidation of this non-toxic substance at ambient temperature is similar to that of TCDD; it can therefore act as a mimic for TCDD in degradative studies. The kinetic investigation was carried out in carbon tetrachloride solution containing DCDD ( $70 \mu\text{g ml}^{-1}$ ) and ruthenium tetroxide ( $4 \times 10^3 \mu\text{g ml}^{-1}$ ); the results given in Table 1 were obtained from the linear plots of log (proportion residual)<sup>-1</sup> to time as illustrated for TCDD.

The rate constants were obtained from the half lives after the induction period from the relation  $k = \ln 2/t_{1/2}$ . The induction period diminished progressively as the reaction temperature rose and in these examples is probably associated with a free radical process, because the formation of radical-cations from dibenzo-*p*-dioxins has been demonstrated<sup>14</sup> and these species have also been detected<sup>15</sup> as intermediates in other oxidations or aromatic compounds by ruthenium tetroxide.

The values for the energy of activation were calculated from the three rate constants using the Arrhenius equation giving a mean value of about 68 kJ mol<sup>-1</sup>. It follows from this that the rate increase per 10 °C rise in this temperature range is about 2.4-fold.

It is possible to show that all members of the PCDD group are susceptible to attack by ruthenium tetroxide. This follows because octachlorodibenzo-*p*-dioxin (OCDD), with the most inactivating chlorine substituents, is itself degraded in refluxing

Table 1 Temperature dependence of the oxidation of DCDD<sup>6</sup> ruthenium tetroxide

| Temperature (°C) | Induction period (min) | Total half life (min) | $t_1$ after induction (min) | $k$ (min <sup>-1</sup> ) | $E_{\text{ACT}}$ (kJ mol <sup>-1</sup> ) |
|------------------|------------------------|-----------------------|-----------------------------|--------------------------|------------------------------------------|
| 30               | 37                     | 215                   | 178                         | $3.89 \times 10^{-3}$    | 66.65                                    |
| 40               | 10                     | 87                    | 77                          | $9.00 \times 10^{-3}$    | 69.1                                     |
| 50               | 4                      | 38                    | 34                          | $2.04 \times 10^{-2}$    |                                          |

carbon tetrachloride ( $C = 70 \mu\text{g ml}^{-1}$ ) by the oxidant ( $4 \times 10^3 \mu\text{g ml}^{-1}$ ), with a half life of ~10 min. This result is also significant because of the possible enzymatic reduction of OCDD to dibenzodioxins of lower chlorine number: analogous changes occur<sup>16</sup> during the metabolism of hexachlorobenzene.

We have no evidence of the nature of the fragments formed during oxidation of the polychlorodibenzodioxins. However, the related chlorophenols undergo extensive decomposition to yield chloride ions and no significant levels of organic products<sup>17</sup>.

The present work shows that ruthenium tetroxide can be used both for the detoxification of glassware and artefacts, and for the periodical purging of industrial reactors to counteract the accumulation of polychloro-*p*-dioxin residues. When using the reagent note that there is some uncertainty in the total reaction time owing to the period of induction. The induction periods given in Table 1 were, however, reasonably reproducible and the delay is not significant at the higher temperature. To obtain information about reaction kinetics the oxidant was used in excess in carbon tetrachloride, but the method is also applicable to the oxidation of PCDDs with catalytic amounts of ruthenium tetroxide as the primary reagent and regeneration by a consumable secondary oxidant, such as hypochlorite.

We thank Dr J. Harrison for DCDD and TCDD and Dr A. J. Dobbs for OCDD. This paper is Part 6 in the series 'Oxidation of Aromatic Substrates'. It was presented in part at the Chemical Society Pollution Symposium in Belfast on 26 September 1979, and subsequently in London on 1 May 1980.

Received 28 October 1980; accepted 27 January 1981.

- Hay, A. *Nature* **281**, 521 (1979).
- Int. Agency Res. Cancer* **15**, 42 (1977).
- Crosland, J. *Ecologist* **10**, 87 (1980).
- Dickens, D. *Nature* **283**, 418 (1980).
- Jansson, B. & Sundström, G. *Sci. Total Envir.* **10**, 209 (1978).
- Dobbs, A. J. & Grant, C. *Nature* **278**, 163 (1979).
- Kearney, P. C., Woolson, E. A. & Ellington, C. P. *Envir. Sci. Technol.* **6**, 1017 (1972).
- Botré, C., Memoli, A. & Alhaique, F. *Envir. Sci. Technol.* **13**, 228 (1979).
- Blair, E. H. (ed.) *Chlorodioxins-Origin and Fate Adv. Chem. Ser.*, No. 120 (American Chemical Society, 1973).
- Cookson, C. *Nature* **278**, 108 (1979).
- Josephson, J. *Envir. Sci. Technol.* **14**, 1165 (1980).
- Lee, D. G. & van den Engh, M. in *Oxidation in Organic Chemistry* (ed. Trahanovsky, W. S.) (Academic, New York, 1973).
- Ayres, D. C. & Scott, C. M. *Envir. Sci. Technol.* **13**, 1983 (1979).
- Shine, H. J. & Shade, L. R. *J. heterocycl. Chem.* **11**, 139 (1974).
- Ayres, D. C. & Gopalan, R. *JCS Chem. Commun.* 890 (1976).
- Mehendale, H. M., Fields, M. & Matthews, H. B. J. *Agric. Food. Chem.* **23**, 261 (1975).
- Ayres, D. C. & Levy, D. P. (in preparation).



## A new pyroxene fractionation trend from a layered basic intrusion

J. R. Wilson\*, K. H. Esbensen† & P. Thy‡

\* Department of Geology, University of Aarhus, 8000 Århus C, Denmark

† Department of Metallurgy, Technical University of Denmark, Building 204, 2800 Lyngby, Denmark

‡ Nordic Volcanological Institute, University of Iceland, 101 Reykjavik, Iceland

**Pyroxene fractionation trends in igneous rocks depend on the magma composition and the conditions of crystallization. Four main types of pyroxene crystallization trend have been recognized in terrestrial igneous rocks (Fig. 1). These are the clinopyroxene trend of mildly alkaline basic magmas<sup>1,2</sup>; the clinopyroxene trend of strongly alkaline basic magmas<sup>3</sup>; the two-pyroxene trend typical of tholeiitic magmas<sup>4,6</sup>; and the two-pyroxene trend typical of calc-alkaline intrusions<sup>7-9</sup>. The calc-alkaline pyroxene trends show restricted iron-enrichment as high  $P_{H_2O}$  causes the crystallization of calc amphibole (and biotite) in place of pyroxenes at a fairly early stage of fractionation. Ca-rich and Ca-poor pyroxenes from the Fongen-Hyllingen complex which coexist with hydrous mafic phases and initially have compositions very similar to those from calc-alkaline complexes, show extreme iron-enrichment in the late differentiates. This is the first recorded example of such extreme pyroxene fractionation in elevated  $P_{H_2O}$  conditions and this fractionation trend, accompanied by other distinct mineralogical features, warrants designation as a new type: the Fongen-Hyllingen trend.**

The 160 km<sup>2</sup> Fongen-Hyllingen complex is an extremely differentiated basic intrusion, synorogenically emplaced in the Scandinavian Caledonides of the Trondheim region, Norway<sup>10</sup>, (11.5°E, 63°N). The complex was emplaced at a pressure of 5–6 kbar before the regional deformation and metamorphism in amphibolite facies<sup>11</sup> peaked. Sufficient primary features remain for igneous trends to be studied in considerable detail.

The complex consists dominantly of layered cumulates, of which a total thickness of ~10 km occurs in the intrusion as a whole. Cryptic layering reveals reversals to higher-temperature mineral compositions with increasing height in the cumulate stratigraphic succession, which is believed to reflect repeated magma influx<sup>12</sup>. Despite these reversals, the highest temperature mineral assemblage in the complex occurs at the base (olivine Fo<sub>86</sub> + picotite) and the lowest temperature assemblage occurs at the top in quartz-bearing ferrosyenites (albite, hedenbergite, ferro-edenite, ilmenite, apatite, quartz, K-feldspar, zircon, biotite and allanite). These syenitic differentiates, which grade downwards into layered ferrogabbroic cumulates, occupy <3% of the exposed area of the complex. Taking the entire stratigraphic sequence, olivine spans the range Fo<sub>86</sub> to Fo<sub>0</sub> with a hiatus between about Fo<sub>71</sub> and Fo<sub>61</sub>, and plagioclase covers the range from An<sub>80</sub> to An<sub>1.5</sub>. Ca-rich pyroxenes vary from Wo<sub>45</sub>En<sub>44</sub>Fs<sub>11</sub> to Wo<sub>47</sub>En<sub>0</sub>Fs<sub>53</sub> and Ca-poor pyroxenes span the range Wo<sub>2</sub>En<sub>67</sub>Fs<sub>31</sub> to Wo<sub>2</sub>En<sub>17</sub>Fs<sub>81</sub>. Inverted pigeonite first occurs at Fs<sub>67</sub> (calculated on a Ca-free basis) and occurs sporadically together with primary orthopyroxene at compositions more iron-rich than this. Apparently magmatic grunerite occurs instead of Ca-poor pyroxenes in the later differentiates and reaches an Mg:Fe ratio of 12:88. Magnetite and ilmenite crystallize over wide temperature intervals. Calcic amphibole is an intercumulus phase throughout most of the complex, becoming a liquidus phase in the upper parts, as well as in sporadic layers in the lower part of the cumulate succession. Its composition ranges from titanian pargasite to ferro-edenite. Biotite is a liquidus phase in the most fractionated rocks.

The coexistence of two pyroxenes, the quartz-bearing nature of the final differentiates and the 'olivine gap' broadly indicate a

tholeiitic parental magma. The detailed igneous petrology of the complex will be published elsewhere.

Broad beam microprobe analyses, ensuring reliable bulk compositions, were made of 30 Ca-rich and 20 Ca-poor pyroxenes representing the entire cumulate succession. Analyses of the most Mg-rich and Fe-rich pyroxenes are given in Table 1. While there is considerable systematic variation in the minor elements Ti, Al, Mn and Na the major trend of crystallization, represented in Fig. 1, is one of increasing Fe:Mg ratio and relatively constant Ca for both Ca-rich and Ca-poor pyroxenes.

The Fongen-Hyllingen Ca-rich pyroxene trend is close to the calcic augite-calcic ferroaugite-hedenbergite trend of the Shiant Isles sill clinopyroxenes<sup>1</sup>, but the latter, which crystallized from a mildly alkaline basalt parent, do not coexist with Ca-poor pyroxenes. Clinopyroxenes of strongly alkaline basic magmas, such as those of the Shonkin Sag laccolith<sup>3</sup>, are enriched in Na and Fe<sup>3+</sup> in the later stages and consequently their evolution, which is not ideally expressed in a En-Di-Hd-Fs diagram, bears no comparison with our pyroxenes. The Fongen-Hyllingen two pyroxene trend is very different in character from the Skærgaard solidus trend<sup>4,5</sup> (augite-ferroaugite-ferrohedenbergite with coexisting bronzite-inverted pigeonite) in the calcic nature of the Ca-rich pyroxenes, the extent of Fe-enrichment of the Ca-poor pyroxenes, and the very late appearance of inverted pigeonite. However, the Fongen-Hyllingen pyroxenes are very close to the Skærgaard subsolidus pyroxene trends<sup>13</sup>. The latter trends are based on microprobe point analyses of individual pyroxene exsolution lamellae/hosts, while our analyses were carried out using a defocused beam (diameter ~20 µm; exsolution lamellae on the micrometre scale) to provide solidus

**Table 1** Broad-beam microprobe analyses of the most Mg and Fe-rich pyroxene compositions observed in the Fongen-Hyllingen complex

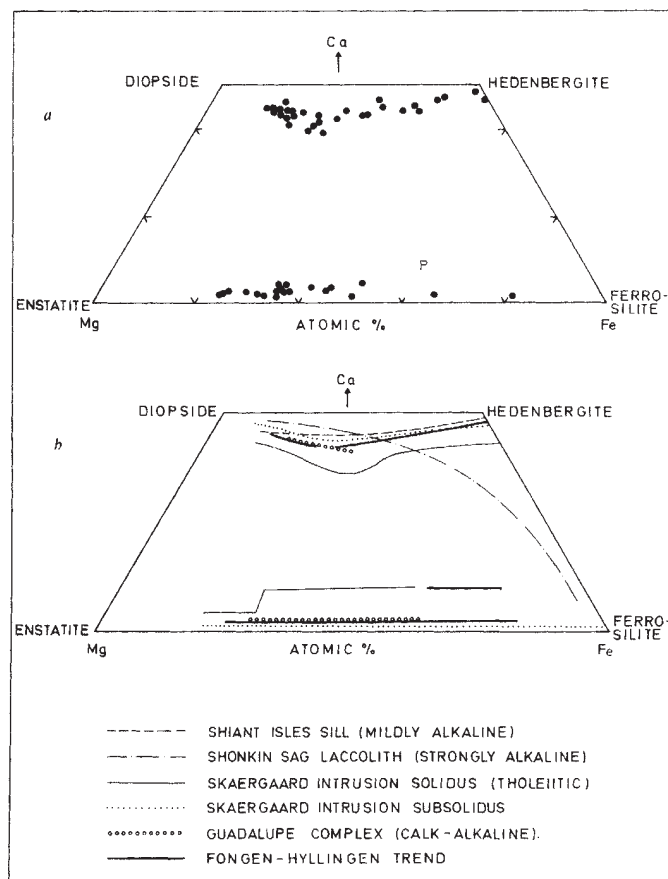
| Analysis no.                   | Ca-rich pyroxene |        | Ca-poor pyroxene |        |
|--------------------------------|------------------|--------|------------------|--------|
|                                | A                | B      | C                | D      |
| SiO <sub>2</sub>               | 49.58            | 48.73  | 53.37            | 47.29  |
| TiO <sub>2</sub>               | 1.03             | †      | 0.26             | 0.11   |
| Al <sub>2</sub> O <sub>3</sub> | 5.07             | 0.26   | 2.97             | 0.42   |
| FeO*                           | 7.10             | 29.75  | 15.33            | 45.40  |
| MnO                            | 0.15             | 0.78   | 0.35             | 1.62   |
| MgO                            | 15.20            | †      | 26.63            | 5.34   |
| CaO                            | 21.55            | 20.63  | 1.13             | 0.90   |
| Na <sub>2</sub> O              | 0.50             | 0.35   | †                | †      |
| Total                          | 100.18           | 100.50 | 100.04           | 101.08 |

\* All Fe expressed as FeO.

† Below detection limit.

compositions. This similarity is, therefore, more apparent than real and the Fongen-Hyllingen pyroxene compositions are consistent with crystallization under elevated  $P_{H_2O}$ , as indicated by the widespread occurrence of primary hydrous mafic phases.

The Fongen-Hyllingen pyroxene compositions form extensions of the calc-alkaline pyroxene trends of the Guadalupe complex<sup>7,8</sup>. Guadalupe Ca-rich pyroxenes are restricted to the augite field and span the range Wo<sub>43</sub>En<sub>45</sub>Fs<sub>12</sub> to Wo<sub>43</sub>En<sub>30</sub>Fs<sub>27</sub>, while the Ca-poor pyroxenes, which show no inversion from pigeonite, extend from Wo<sub>3</sub>En<sub>69</sub>Fs<sub>28</sub> to Wo<sub>4</sub>En<sub>35</sub>Fs<sub>61</sub>. As a result of crystallization under fairly high  $P_{H_2O}$  in calc-alkaline complexes, the pyroxene liquidus surface is depressed, producing Ca-rich clinopyroxenes and extending the field of primary orthopyroxene crystallization to iron-rich compositions. Elevated  $P_{H_2O}$  also results in calcic amphibole being a liquidus phase at an early stage of fractionation, crystallizing together with two pyroxenes until the anhydrous pyroxenes become unstable and first Ca-rich pyroxene and then Ca-poor pyroxene cease to crystallize<sup>14</sup>. Calcic amphibole, together with biotite, then crystallize as the only mafic minerals until the end of fractionation.



**Fig. 1** *a*, Pyroxene compositions from the Fongen-Hyllingen complex. *P*, first occurrence of inverted pigeonite. No inverted pigeonites have been analysed. *P* is located in the pyroxene quadrilateral on tie line between coexisting calcic ferroaugite and ferrohypsthene. *b*, Comparison of the Fongen-Hyllingen pyroxene crystallization trend with those of other igneous complexes.

The extreme pyroxene trend presented here, with highly calcic Ca-rich pyroxene reaching Mg-free hedenbergite and a very extensive Ca-poor pyroxene trend with two Ca-poor pyroxenes coexisting at Fe-rich compositions, represents crystallization under moderately elevated and increasing  $P_{H_2O}$ . Water vapour pressure was higher than that typical of tholeiitic complexes but not high enough to produce the more restricted calc-alkaline trend. Other mineralogical features, such as the wide plagioclase and olivine fractionation ranges (reaching albite and pure fayalite respectively) and the occurrence of a small volume of syenitic differentiates as opposed to large volumes of granitic material, also distinguish the Fongen-Hyllingen fractionation from calc-alkaline complexes. We propose that this new pyroxene fractionation trend be named the Fongen-Hyllingen trend.

The Danish Natural Science Research Council and the Carlsberg Foundation provided financial support to study the Fongen-Hyllingen complex.

Received 17 November 1980; accepted 21 January 1981.

- Gibb, F. G. F. *J. Petrol.* **14**, 203 (1973).
- Aoki, K. *Am. Miner.* **49**, 1199 (1964).
- Nash, W. P. & Wilkinson, J. F. G. *Contr. Miner. Petrol.* **25**, 241 (1970).
- Brown, G. M. *Miner. Mag.* **31**, 511 (1957).
- Brown, G. M. & Vincent, E. A. *J. Petrol.* **4**, 175 (1963).
- Atkins, F. B. *J. Petrol.* **10**, 222 (1969).
- Best, M. G. *J. Petrol.* **4**, 223 (1963).
- Best, M. G. & Mercy, E. L. P. *Am. Miner.* **52**, 436 (1967).
- Nockolds, S. R. *J. geol. Soc. Lond.* **96**, 451 (1941).
- Wilson, J. R. & Olesen, N. Ø. *Norsk geol. Tidsskr.* **55**, 423 (1975).
- Olesen, N. Ø., Hansen, E. S., Kristensen, L. H. & Thyrsed, T. *Leid. geol. Meded.* **49**, 259 (1973).
- Esbensen, K. H., Thy, P. & Wilson, J. R. *Norsk geol. Tidsskr.* **58**, 103 (1978).
- Nwe, Y. Y. *Contr. Miner. Petrol.* **55**, 105 (1976).
- Cawthorn, R. G. *Am. Miner.* **61**, 907 (1976).

## Submarine chutes on the slopes of fjord deltas

David B. Prior & Wm J. Wiseman Jr

Coastal Studies Institute, Louisiana State University, Baton Rouge, Louisiana 70803, USA

W. R. Bryant

Department of Oceanography, Texas A and M University, College Station, Texas 77843, USA

The first side-scan sonograph images of submarine landslide chutes from arctic fjords presented here indicate that submarine slopes of fjord head deltas are subject to instability. The chutes transport sediment from the upper delta slopes into deeper water parts of the fjords. Previously, sediment distribution and budgets in the bottom of deepwater fjords were ascribed to settling of suspended sediment supplied by streams and rivers during the short summer melt period. Previous descriptions of submarine mass movements from some subarctic fjords have been limited to isolated major catastrophic events. The sonographs from an arctic fjord suggest that delta-front instability occurs as multiple small-scale events that recur at frequent intervals and are intrinsic to the fjord sedimentation system.

The underwater topography of Adventfjord, a Spitsbergen fjord, and in particular actively prograding fluvio-glacial deltas and their prodelta slopes, was examined using a side-scan sonar system. Chutes or delta-front gullies were distinctive features in some locations. Although the precise origins and functions of the chutes are obscure, morphological elements suggest that they may be due to underwater landsliding or mass movement processes; however, they are undoubtedly a significant component of some fjord delta geometries and associated progradational and depositional mechanisms. The present sonographs are from a small fan delta formed by active deposition from Longyear River into the side of Adventfjord and a major delta at the head of the fjord, the result of deposition from a braided outwash stream network, Advent River.

The submarine slopes in the fjord are steep. Water depths reach 30 m very close to the shoreline. Each delta has a well-developed planar bar formed by coalescence of adjacent distributary mouth bars, which produces a relatively flat feature composed of sand, silt, and fine gravel and covered by 1–3 m of water at low tide (tidal range 1–2 m). The delta-front slopes, seawards of the bars, are steeply inclined at angles of 12–16°. At the base of the delta slopes the fjord bottom is essentially flat at water depths of 60 m. The deltas demonstrate classical slope geometry and sediment lithology with fine sands on the bars and upper slopes and grain size becoming progressively finer seaward.

Chutes were observed in the upper slopes of the deltas between 5 and 30-m water depth, that is, where bottom slope angles are greatest. Typically the chutes begin on the steepest parts of the delta on the seaward margin of constructional aprons or bars, and are not incised into the nearshore surfaces of the bars. Thus, the chutes are not simply seaward extensions of distributary channels over the top of the submerged bars, but rather originate in arcuate re-entrants, with steep head slopes, cut into the bar fronts. The chutes usually exhibit an upslope bowl-shaped source region with an upslope scarp, an elongated narrow gully section, and a downslope extensive hummocky zone.

Figure 1 shows a sonograph and interpretive map of a system of chutes radiating out from the bar front of the Longyear fan delta. The bar is extremely irregular, with some large scarped and stepped areas. The scarps occasionally seem to bound



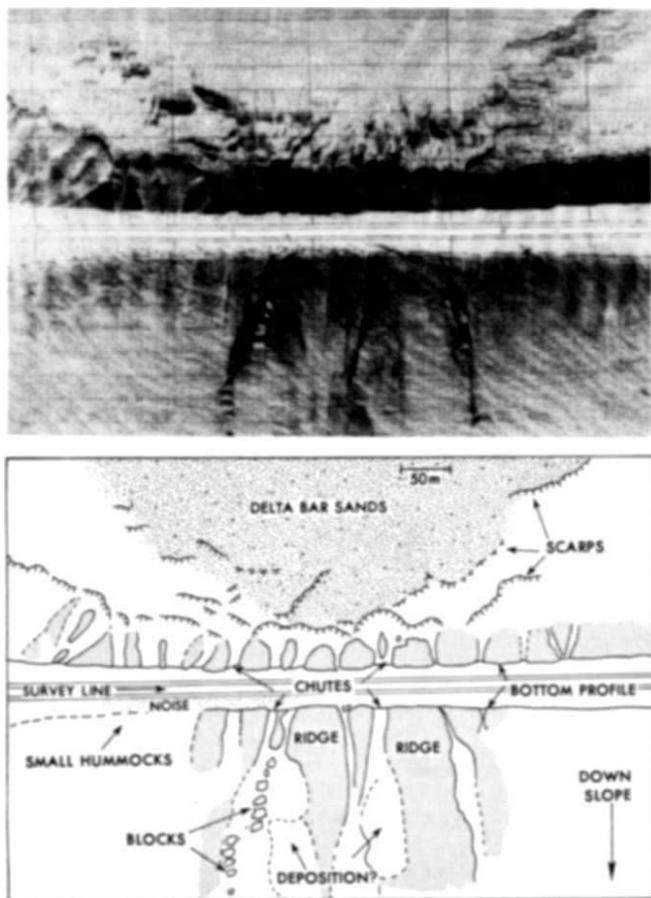


Fig. 1 Side-scan sonograph of Longyear delta front and interpretive map.

masses that have subsided and moved seawards. A fairly continuous scarp, extending around the front of the delta, shows major re-entrants immediately above distinct chutes. The chutes begin as bowl-shaped subsidence areas but become narrower farther downslope. The subsidence areas are typically 20–25 m across, whereas the elongate chutes average 10 m in width, are incised into the delta slope to depths of 2–5 m, and have no identifiable associated levees. The chutes show two distinct tonal patterns on this sonograph. In the central part of the delta, five chutes show very dark grey tones, whereas residual interchute ridges have relatively lighter grey tones. In contrast, chutes towards the edges of the fan are lighter toned than the intervening ridges. The precise reasons for such striking tonality differences are unknown but may be due to relative age, the darker toned chutes being relatively younger with incision exposing coarser-textured delta-front deposits, the lighter toned chutes being partially infilled by fine-grained deposits settling from suspension. Figure 1 also reveals that the lower (downslope) ends of the five freshest features are marked by irregularly shaped depositional areas of blocks and spreads of dark-toned sediments.

Figure 2 shows a series of chutes on the braided outwash delta at the head of Adventfjord. Chutes again emanate from bowl-shaped source areas backed by scarps on the edge of the distal bar. The chutes are represented by lighter toned areas separated by dark-toned interchute ridges and intact platforms. They vary in width from 15 to 30 m. In some cases adjacent chutes coalesce. The incisions are relatively shallow (<3 m) where the survey line crosses the chutes. Downslope from the chutes the delta-front slope is very hummocky, with contrasting light and dark tones, indicating considerable local relief over a large area. At one place, where several chutes converge, the downslope

area consists of several subparallel blocks and scarps. Other survey data show that the chutes are present across the entire bar front and that the middle delta slope is hummocky. The hummocks, blocks, and scarps decrease in size towards the foot of the delta slope, finally merging with the flat fjord floor.

The origin of the features in Figs 1 and 2 is not precisely known. However, an attractive hypothesis is that the bowl sources, chutes, and hummocky slopes are due to submarine sediment mass movement on the actively prograding delta slopes. Mass movement (slope instability) is an important process in many other active delta environments. For example, detailed sonar mapping in the Mississippi Delta region has revealed similar mudslide gullies<sup>1–6</sup>. Other examples include the Copper River delta (Alaska)<sup>7</sup>, the Fraser Delta (British Columbia)<sup>8</sup>, and the Esmeraldas and Magdalena deltas (South America)<sup>9,10</sup>. There is increasing evidence of sea floor sediment mass movement specifically in subarctic fjords and glacier-fed lacustrine delta areas. For example, one classic fan delta-front failure in Howe Sound fjord (British Columbia) caused extensive damage to adjacent buildings<sup>11</sup>, as did slide-generated waves in Kitimat Inlet (British Columbia)<sup>12</sup>. A survey of the Howe Sound delta with side-scan sonar showed the presence of delta-front chutes<sup>13</sup>. Similarly, the geological interpretation of relict and ancient fjord deltaic sediments is relying more and more on mass movement concepts and processes<sup>14</sup>.

The combination of circumstances in fjord deltas are conducive to slope failure. Bottom slopes are steeply inclined, and rapid progradation, especially during spring melt, causes temporary oversteepening and increases gravitational stress. Rapid deposition of sands, silts, and clays can cause local excess pore pressures, especially where fine sediment layers impede pore fluid drainage. Furthermore, bar sands are built out over previously deposited fjord bottom clays and glacial flour; the latter may not be able to accept the prograding bar loads, and

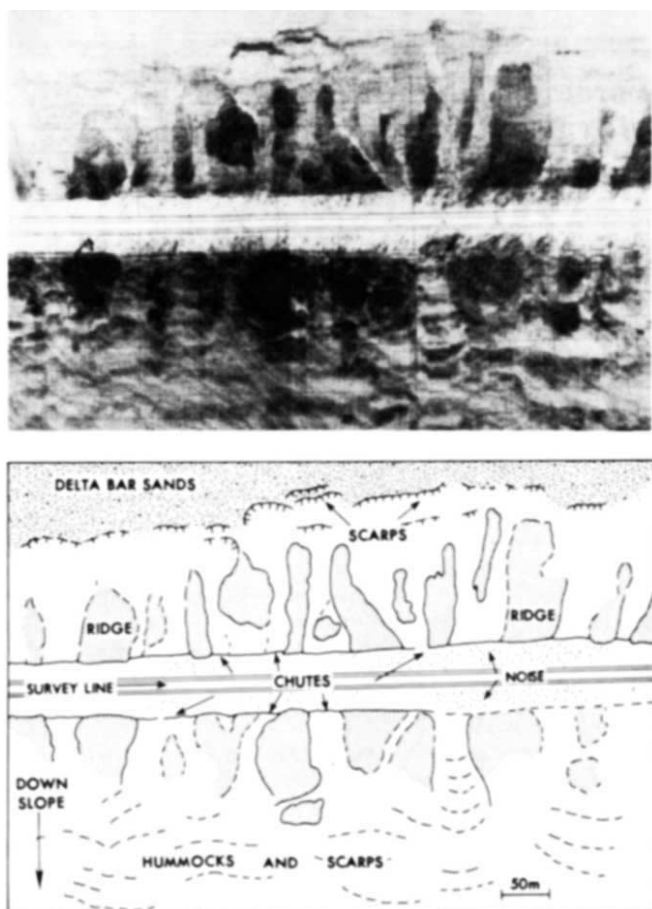


Fig. 2 Side-scan sonograph of Adventfjord delta front and interpretive map.

failures may result. Finally, upfjord winds can cause surface waves, which, as they shoal over the edge of the bar, may impart significant cyclic loading effects on the bar-front sediments.

The chutes and subsidence areas on delta slopes in Spitsbergen can be interpreted as the result of avalanching or debris flow movement of the bar-front sediments. A cyclic wave-induced loading of the sediments can generate pore fluid pressures sufficient to liquefy the bar-front deposits, which are already at or near the submerged angle of repose due to normal rapid sedimentation. Similarly, locally rapid depositional oversteepening could have the same effect. The transport of sediments down the chutes transfers loading stresses to the middle slopes of the delta, which may then, in turn, suffer failure, producing irregular hummocky terrain.

As similar side-scan sonar surveys are made for other fjord deltas, it is anticipated that delta-front chutes and other results of mass movement processes may prove to be extremely common.

This study was sponsored by the Arctic Program and Coastal Science Program of the Office of Naval Research, Arlington, Virginia.

Received 14 October 1980; accepted 23 January 1981.

1. Prior, D. B. & Coleman, J. M. *Geoscience and Man* Vol. 19 41–53 (Louisiana State University, Baton Rouge, 1978).
2. Prior, D. B. & Coleman, J. M. *Mar. Geotech.* **3**, 37–60 (1978).
3. Prior, D. B., Coleman, J. M. & Garrison, L. E. *Geology* **7**, 423–425 (1979).
4. Prior, D. B., Coleman, J. M., Suhayda, J. N. & Garrison, L. E. *Proc. Conf. on Port and Ocean Engineering under Arctic Conditions*, Trondheim, **2**, 921–933 (1979).
5. Prior, D. B. & Coleman, J. M. *Mar. Geol.* **36**, 227–239 (1980).
6. Coleman, J. M., Prior, D. B. & Garrison, L. E. *Bureau of Land Management Rep.* 80–81 (New Orleans, 1980).
7. Hampton, M. A. *et al. Offshore Tech. Conf.* Houston, Texas, Pap. 3314, 2307–2318 (1978).
8. Mathews, W. H. & Shepard, F. P. *Bull. Am. Ass. petrol. Geol.* **46**, 1416–1437 (1962).
9. Shepard, F. P. *Trans. Am. geophys. Un.* **13**, 226–230 (1932).
10. Heezen, B. C. *Bol. Soc. Geogr. Colombia* **51–52**, 135–143 (1956).
11. Terzaghi, K. *Proc. 8th Texas Oil Mech. Engr. Conf.* 1–41 (1956).
12. Murty, T. S. *J. geophys. Res.* **84** (C12), 7777–7779 (1979).
13. Prior, D. B., Wiseman, Wm. J. & Gilbert, R. (in preparation).
14. Shaw, J. *Geogr. Annlr* **59** (A), 221–240 (1977).

## Foraminifera distribution of provinces in the Gulf of Mexico

Stephen J. Culver & Martin A. Buzas

Department of Paleobiology, Smithsonian Institution, Washington DC 20560, USA

**Recent benthic foraminifera distribution patterns form the basis of much palaeoenvironmental (particularly palaeobathymetric) interpretation used by geologists. We have recently summarized the distribution of recent benthic foraminifera on the Atlantic continental margin of North America<sup>1,2</sup> and report here the results of an analysis of all published data from the Gulf of Mexico. To obtain a synthesis of foraminiferal distribution in the Gulf of Mexico, a computerized catalogue of all published occurrences (presence or absence data) was compiled. Cluster analysis of these data isolated four large, marginally overlapping areas or provinces exhibiting a spatial correlation with Gulf of Mexico water masses. A fifth, smaller, discrete area (biofacies) was found at the mouth of the Mississippi delta.**

There have been 77 papers in the past 60 years describing the distribution of recent benthic foraminifera in the Gulf of Mexico. We compiled this information using a large computer system. Our files record the distribution of 1,219 species names from 426 localities<sup>3</sup>. Through synonymization using published illustrations and specimens deposited in the US National Museum and the British Museum, the number of species was reduced to 848. Because cluster analysis of commonly occurring species on the Atlantic continental margin produced comparable results to analysis involving all species<sup>4</sup>, we used the 295 most commonly recorded species (those occurring at 4% or more of the 426 localities) for cluster analysis of the Gulf of Mexico data. Due to computer limitations the number of locali-

ties involved in analysis was reduced randomly to 350. The data matrix was analysed in the *Q*-mode by clustering Jaccard coefficients by the weighted pair-group method<sup>5–7</sup>.

The cluster analysis produced four large, marginally overlapping areas or provinces (Fig. 1): the Coastal, Inner Shelf, Outer Shelf, and Slope and Abyssal Plain Provinces. (The diagnostic species associations of the provinces will be published elsewhere.) A fifth, smaller, discrete area or biofacies is located at the mouth of the Mississippi River.

Our analysis, based on all published presence or absence data for 295 commonly recorded species of benthic foraminifera, indicate that for the sampled littoral, inner shelf and outer shelf areas, the same foraminifera assemblages occur all around the Gulf and comprise the Gulf of Mexico Coastal, Inner Shelf and Outer Shelf Provinces respectively (Fig. 1). In deeper water, the Gulf of Mexico Slope and Abyssal Plain Province covers the whole sampled sea bed beneath depths of ~1,000 m (Fig. 1). Our circum-Gulf provinces agree with the pattern of benthic foraminifera distributions reported by Poag<sup>8</sup>, but he divided the deep basin into distinct eastern and western assemblages as did Dignes<sup>20</sup>.

The provinces defined by our analysis seem to be unrelated to sediment substrate. For example, the Yucatan and west Florida carbonate shelves and the clastic northern Gulf of Mexico shelf east and west of the Mississippi delta are all included in the Gulf Inner Shelf Province. In agreement with macrofaunal patterns<sup>9</sup>, therefore, foraminifera shelf faunas on both sides of the Mississippi delta are generally the same. However, at the mouth of the delta, 10 localities group together as a discrete biofacies (Fig. 1). This group is considered to be too small and geographically restricted to be referred to as a province, but its low diversity assemblage shows close affinities to the Outer Shelf Province assemblage. The position of this biofacies is probably related to oceanographic or sedimentological conditions related to the Mississippi River discharge.

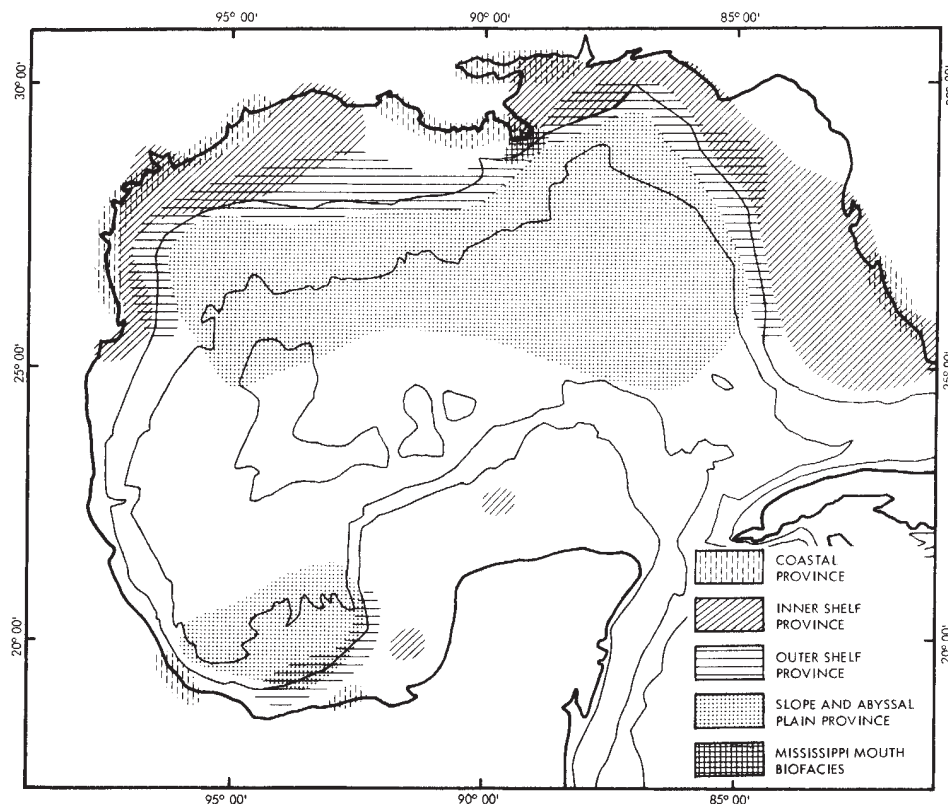
Various provincial schemes have been proposed for the Gulf of Mexico–Caribbean region using shallow water macro-invertebrates and fishes<sup>9–15</sup>. The northern Gulf coast from Cape Romano, Florida to approximately the Mexico–USA border is now generally considered to be part of the east coast of North America Carolinian Province<sup>9,14,15</sup>. The southern Gulf coast and the southern tip of Florida from Cape Romano to Cape Canaveral is part of the Caribbean Province, which includes the West Indian islands and the northern coast of South America<sup>9,14,15</sup>. Other schemes involve greater subdivision<sup>10,13</sup> or the inclusion of all the Gulf of Mexico into a Caribbean Province<sup>12</sup>.

We compared the shallow-water benthic foraminifera assemblages from the east coast of North America south of Cape Hatteras with shallow water assemblages in the Gulf of Mexico (data unpublished). An inner continental shelf Carolinian Province, similar in extent to that defined by macro-invertebrate assemblages (see above), does not exist for benthic foraminifera. That is, inner shelf foraminifera faunas of the northern Gulf of Mexico differ considerably from those of the shelf south of Cape Hatteras. Sixty commonly recorded species are present on the east coast Southern Shelf Province, and 172 are present in the Gulf of Mexico Inner Shelf Province; only 42 species (22% of a total of 189 species) are commonly recorded species in both provinces. The question remains whether the Gulf of Mexico, based on foraminifera data, is part of a larger Caribbean Province.

A congruous relationship between foraminifera provinces and bottom water masses in the Gulf of Mexico is demonstrable. As for the Atlantic Coastal Province<sup>1</sup>, we suggest the Gulf of Mexico Coastal Province is related to the marginal waters of bays, estuaries and lagoons. The Inner and Outer Shelf Provinces show varying amounts of overlap. The distinction between these provinces can be related to the difference in magnitude of seasonal effects between the inner and outer shelf. The critical limits of the set of environmental variables affecting the foraminifera assemblages will be located at different



**Fig. 1** Recent benthic foraminifera provinces in the Gulf of Mexico; 183 m, 1,830 m, 3,660 m submarine contours are shown.



distances from the shore from year to year (or season to season): hence the considerable degree of overlap. The Inner Shelf Province on the southern margin of the Gulf may be less well defined due to sampling deficiencies or to decreased environmental variability, in particular less severe salinity changes due to a relative lack of river discharge compared with the northern margin of the Gulf.

The depth limit of seasonal effects on water mass characteristics is ~100–200 m (refs 21, 22). This depth correlates with the basinward limit of the Inner Shelf Province. Although the open surface waters of the Gulf can be divided into eastern and western parts<sup>16</sup>, the deeper water, below ~1,000–1,500 m, has generally uniform temperature, salinity and dissolved oxygen relations<sup>16–19</sup>. Thus, the disposition of the pan-Gulf of Mexico Slope and Abyssal Plain Province (Fig. 1) agrees with that of the bottom water mass (that is, no division into eastern and western provinces) and an upper limit at ~1,000 m. In agreement with the provincial pattern presented here, distribution maps for Gulf of Mexico commonly recorded species<sup>3</sup> show that all species, except one, that occur at abyssal depths range up on to the slope or even the outer shelf. Thus, there is no solely abyssal association composed of species restricted to abyssal depths.

We thank C. G. Adams, L. Isham, A. Lanham, C. McCloy, K. McCormick, D. Piacesi, D. Schnitker, D. J. Stanley, R. Cifelli and J. Whittaker for assistance. The research was supported by a Smithsonian Scholarly Studies Award.

Received 11 December 1980; accepted 3 February 1981.

1. Buzas, M. A. & Culver, S. J. *Science* **209**, 687–689 (1980).
2. Culver, S. J. & Buzas, M. A. *Smithson. Contr. Mar. Sci.* **6**, 1–512 (1980).
3. Culver, S. J. & Buzas, M. A. *Smithson. Contr. Mar. Sci.* (in the press).
4. Culver, S. J. & Buzas, M. A. *J. For. Res.* (in the press).
5. Sokal, R. R. & Sneath, P. H. A. *Principles of Numerical Taxonomy* (Freeman, San Francisco, 1963).
6. Kaesler, R. L. *Kans. Paleontol. Contr.* **10**, 1–50 (1966).
7. Mello, J. F. & Buzas, M. A. *J. Paleontol.* **42**, 747–758 (1968).
8. Poag, C. W. *Abstr. Benthonics '75 '35* (Dalhousie University, Nova Scotia, 1975).
9. Hedgpeth, J. W. *Inst. Mar. Sci., Univ. Texas* **3**, 111–224 (1953); *Fish. Bull.* **89**, *Fish Wildlife Serv.* **55**, 203–214 (1954).
10. Briggs, J. C. *Marine Zoogeography* (McGraw-Hill, New York, 1974).
11. Fischer, P. *Manuel de Conchyliologie* (Savy, Paris, 1887).
12. Hall, C. A. *Ecology* **45**, 226–234 (1964).
13. Pulley, T. E. A. *Rep., Am. Malacol. Un.* 1952, 2–3 (1953).
14. Rehder, H. A. *Fish. Bull.* **89**, *Fish Wildlife Serv.* **55**, 469–474 (1954).
15. Warmke, G. L. & Abbott, R. T. *Caribbean Seashells* (Livingston, Narberth, 1961).
16. Ichiye, T. *Geofis. Int.* **2**, 47–76 (1962).

17. Harding, J. L. & Nowlin, W. D. Jr *Encyclopedia of Oceanography* (ed. Fairbridge, R. W.) 324–331 (Rheinhold, New York, 1966).
18. McLellan H. J. & Nowlin, W. D. Jr *J. Mar. Res.* **21**, 233–245 (1963).
19. Nowlin, W. D. Jr *Contributions to the Physical Oceanography of the Gulf of Mexico* (eds Capurro, L. R. A. & Reid, J. L.) 3–51 (Gulf, Houston, 1972).
20. Dignes, T. W. thesis, Univ. Maine (1978).
21. Nowlin, W. D. Jr & Parker, C. A. *J. phys. Oceanogr.* **4**, 467–486 (1974).

## Mimicry of antenna and photo-protective carotenoid functions by a synthetic carotenoporphyrin

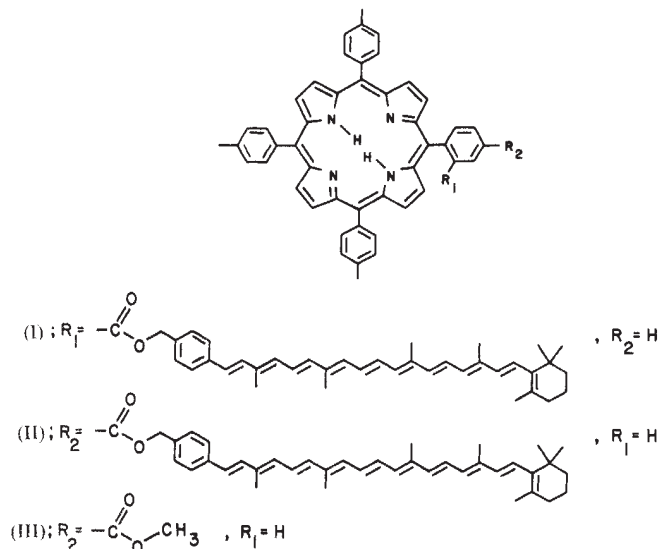
René V. Bensasson\*, Edward J. Land†, Ana L. Moore‡, Robert L. Crouch‡, Gary Dirks‡, Thomas A. Moore‡ & Devens Gust‡

\* Laboratoire de Biophysique, INSERM U201, Muséum National d'Histoire Naturelle, 61 Rue Buffon, 75005 Paris, France

† Paterson Laboratories, Christie Hospital and Holt Radium Institute, Manchester M20 9BX, UK

‡ Department of Chemistry, Arizona State University, Tempe, Arizona 85281, USA

Carotenoid pigments, ubiquitous in photosynthetic membranes, are essential for the survival of green plants<sup>1</sup>. Two facets of carotenoid function are recognized in photosynthetic membranes. First, carotenoids prevent the chlorophyll-photo-sensitized formation of highly destructive singlet oxygen by intercepting the chlorophyll triplet states<sup>2–10</sup> and may also scavenge additional singlet oxygen present<sup>11,12</sup>. Second, carotenoids perform an antenna function by transferring the energy of absorbed light at the singlet excited state level to the chlorophyll system for the execution of photosynthetic work<sup>13–16</sup>. Nevertheless, the mechanisms by which carotenoids perform these functions are poorly understood. We now report that a unique synthetic carotenoporphyrin I consisting of a carotenoid part covalently linked to a synthetic tetraarylporphyrin successfully mimics both the photophysical functions of carotenoids in photosynthesis. The explanation for this seems to be the close interaction of the carotenoid and porphyrin  $\pi$ -electron systems.

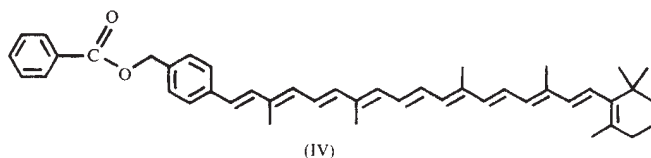


We have previously shown that singlet-singlet energy transfer from the carotenoid moiety of carotenoporphyrin II to the porphyrin is vanishingly small<sup>17</sup>. Thus, simply placing a carotenoid in the proximity of a porphyrin is not enough to ensure antenna function. In I, on the other hand, carotenoid to porphyrin singlet-singlet energy transfer occurs with 25% efficiency<sup>18</sup>. The enhanced energy transfer efficiency in I is due to the fact that unlike II, I adopts a folded conformation in which the carotenoid  $\pi$ -electron system lies directly over that of the porphyrin (Fig. 1). The close approach of the two chromophores suggests that an electron exchange interaction may contribute to the observed singlet-singlet energy transfer<sup>19</sup>. In view of the fact that triplet-triplet energy transfer is necessarily mediated through exchange terms<sup>20</sup>, and because singlet oxygen arises from the interaction of oxygen with a sensitizer triplet<sup>12</sup>, we have investigated the photochemical and dynamic properties of the triplet states of I and II.

Singlet oxygen can be conveniently detected by spectroscopically monitoring its reaction with diphenylisobenzofuran (DPBF)<sup>21</sup>. In Fig. 2, singlet oxygen production detected in this manner is plotted for a variety of sensitizers including I, II, and a simple tetraarylporphyrin, III. In addition, the effect of added

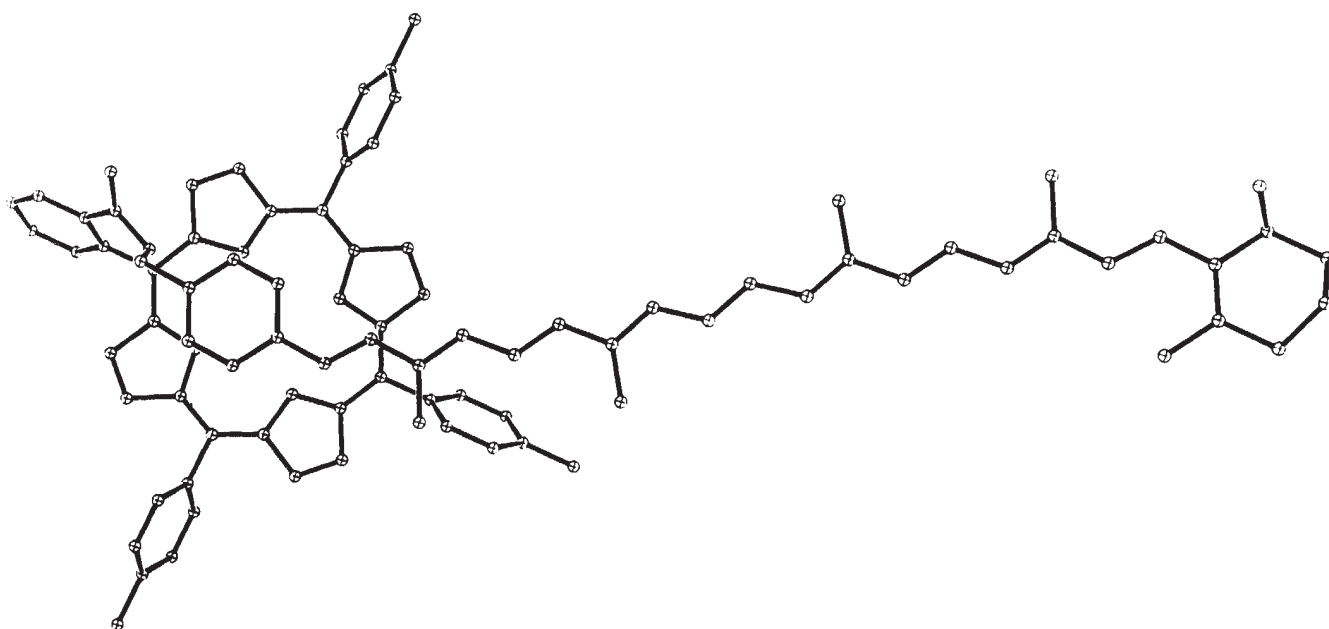
carotenoids on photoproduction of singlet oxygen by III is shown. The most striking feature is the essentially complete photoprotection afforded by I when contrasted with the extensive singlet oxygen production by II. If efficient photoprotection were simply a matter of relative concentration with no specific geometric requirements, one might expect similar degrees of photoprotection in I and II. This is most definitely not the case, and an understanding of the results required a more detailed investigation of the photophysics of these systems.

Consequently, we have carried out pulse radiolysis and laser flash experiments to monitor directly the evolution of the triplet states<sup>22</sup>. The properties of the carotenoid triplet states were established by pulse radiolysis. Population of the lowest triplet states of I, II and IV by pulse radiolysis in benzene through sensitization by biphenyl<sup>22</sup> resulted in a transient spectrum with  $\lambda_{\text{max}} = 535 \text{ nm}$  and  $\epsilon_T - \epsilon_G \sim 1.4 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$ , which is characteristic of a carotenoid triplet<sup>23,24</sup> ( $\epsilon_T$  and  $\epsilon_G$  are the extinction coefficients of triplet and ground states, respectively). Similar experiments performed with III resulted in an  $\epsilon_T - \epsilon_G \sim 6.8 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$  at 440 nm, the transient difference spectrum being typical of a porphyrin. It is important to note that these transient spectra observed through pulse radiolysis for I, II and



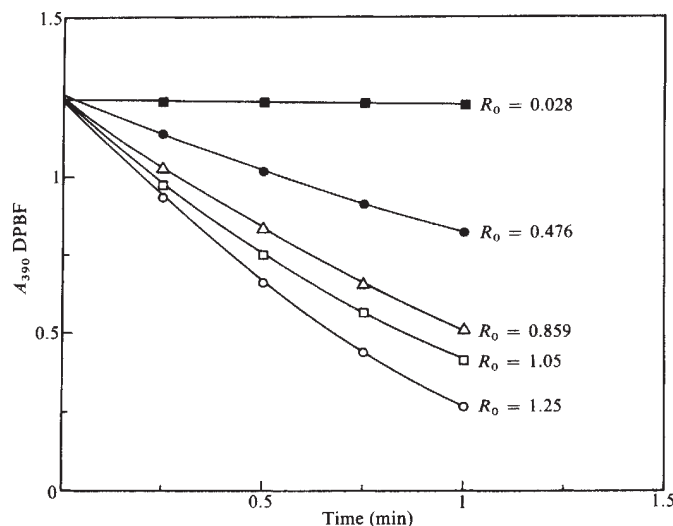
IV are the same: the spectra and extinction coefficients for I and II are not perturbed by the covalent linkage between the two chromophores.

We now turn to the laser flash experiments. Figure 3 presents oscilloscope traces of the transient absorption changes with time for II after 353 nm laser excitation. The porphyrin triplet observed at 440 nm (Fig. 3b) appears first with  $t_{\frac{1}{2}} \sim 1.5 \mu\text{s}$ . Concomitant with this decay, and in a 1:1 ratio, the carotenoid triplet observed at 535 nm (Fig. 3a) rises with the same  $t_{\frac{1}{2}}$ . Subsequently, this absorption decays with  $t_{\frac{1}{2}} \sim 7 \mu\text{s}$  (Fig. 3c). Taking into account the carotenoporphyrin concentration and the maximum possible expected intermolecular triplet energy transfer rate, this result suggests that the carotenoid is quenching the porphyrin triplet state by an intramolecular energy



**Fig. 1** Idealized time average solution conformation for carotenoporphyrin I. The aromatic ring of the carotenoid moiety is  $\sim 4 \text{ \AA}$  above the mean porphyrin plane. Hydrogen atoms are not shown. This conformation is based on proton NMR spectral data<sup>18</sup>.



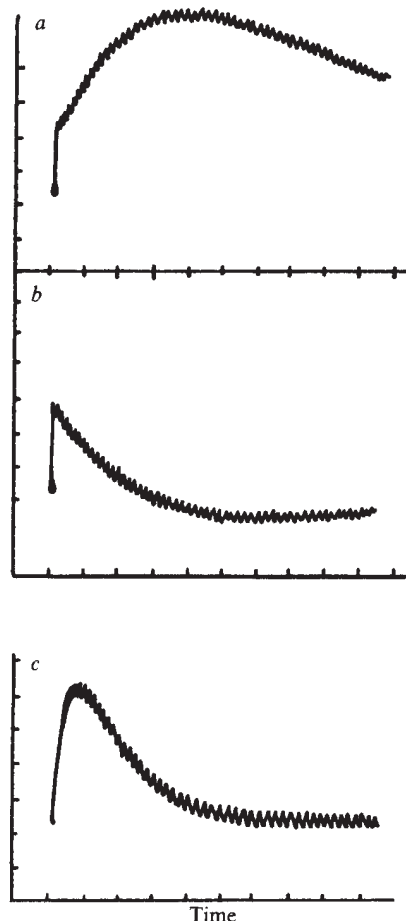


**Fig. 2** Bleaching of DPBF at 390 nm by singlet molecular oxygen. Filtered light (630 nm long pass cutoff, Oriel G-772-6300 and a water filter) from a 150-W tungsten halogen projection lamp was used to illuminate the solutions. Reactions were carried out in spectrophotometer cells open to the air at solution volumes of ~2 ml. Sample temperature was  $23 \pm 1^\circ\text{C}$ . All solutions were prepared from freshly distilled toluene stored over molecular sieves. All compounds were purified by TLC immediately before use except DPBF which was used as received (Sigma). The DPBF concentration was  $2.1 \times 10^{-6}$  M in all experiments.  $\blacksquare$ ,  $1.9 \times 10^{-6}$  M I;  $\bullet$ ,  $1.8 \times 10^{-6}$  M II;  $\triangle$ ,  $1.9 \times 10^{-6}$  M III plus  $1.9 \times 10^{-6}$  M  $\beta$ -carotene;  $\square$ ,  $1.9 \times 10^{-6}$  M III plus  $1.9 \times 10^{-6}$  M benzoate ester of 7'-apo-7'-(4-hydroxymethylphenyl)- $\beta$ -carotene (IV);  $\circ$ ,  $2.0 \times 10^{-6}$  M III. The wavelength monitored (390 nm) was chosen for its relative freedom from absorbance by chromophores other than DPBF. At each point the entire absorption spectrum (200–800 nm) was acquired (Hewlett-Packard 8450 A spectrometer) to ensure that no other photochemical processes were occurring. Plotted absorbance values have been corrected by subtracting residual absorbance at 390 nm of the solution without DPBF (typically  $A_{390}$  0.06–0.09). The initial rates ( $R_0$ ) are  $(A_{t_0} - A_{t_{15}})/15$  s. Illumination of  $2.1 \times 10^{-6}$  M DPBF showed the same response as I within experimental error. There was no change in absorbance at 390 nm when  $2.0 \times 10^{-6}$  M I was illuminated.

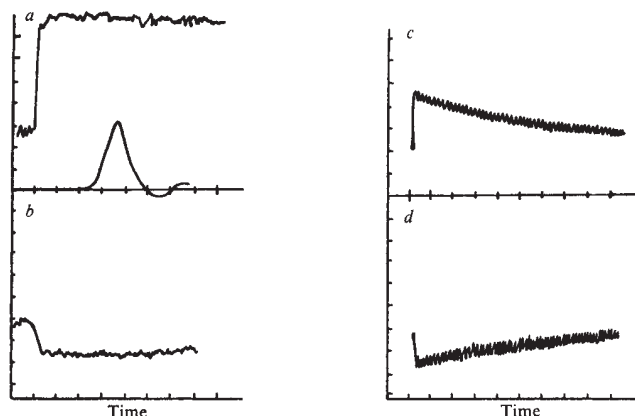
transfer process. Figure 4 presents the results for a similar experiment for I. The behaviour of this carotenoporphyrin is strikingly different. The carotenoid triplet appears immediately (Fig. 3a); the porphyrin triplet state cannot be observed even at the time resolution limit (~30 ns) of our instrument (Fig. 3b). The quantum yield, determined by the comparative method<sup>22</sup>, for the formation of the porphyrin triplet state in III was 0.45 whereas the quantum yield of carotenoid triplet state, and consequently also the yield for the formation of the porphyrin triplet state, was 0.23 in II and 0.12 in I. (These quantum yields are consistent with the increase in quenching of the porphyrin fluorescence from III to II to I<sup>17,18</sup>. The competing nonradiative path from the porphyrin singlet state is of interest because electron transfer between the chromophores may be occurring. This mechanism has been proposed for the quenching of chlorophyll *a* fluorescence by  $\beta$ -carotene<sup>25</sup>.)

From these results it is clear that the remarkable photoprotective function observed for I (see Fig. 2) is a consequence of the extremely rapid quenching of the porphyrin triplet state by the carotenoid moiety. *In vitro* studies of bacterial systems have shown that quenching of the bacteriochlorophyll triplet state by carotenoid occurs with  $t_{1/2} \sim 20$  ns (ref. 6). The difference in porphyrin triplet quenching between I and II shows that achieving the required quenching rates on the nanosecond time scale requires a conformation such as that found for I (Fig. 1), which allows a high degree of orbital overlap and therefore very strong exchange interactions. Such a conformation is not available for II.

It is interesting to speculate on the implications of the above results for the elucidation of the photoprotective and antenna functions of carotenoids *in vivo*. A stacked conformation with close approach of the  $\pi$ -electron systems was necessary in I both for antenna function in the singlet manifold (where energy transfer must compete with rapid internal conversion in the carotenoid)<sup>17,18,27</sup> and for efficient suppression of singlet oxygen production via the triplet manifold (where porphyrin triplet quenching by the carotenoid must compete with quenching by abundant oxygen). Such a conformation was necessary because both processes require relatively strong electronic interactions between the carotenoid and porphyrin  $\pi$ -electron systems. Perhaps a similar arrangement of pigments, with a separation of only a few ångströms, is necessary in photosynthetic reaction centres. Indeed, results for I and II suggest that the photoprotective and antenna functions of a particular carotenoid molecule *in vivo* may involve the same chlorophyll molecule. Such intimate and specific pigment-pigment interactions are doubtless mediated by the higher-order structure of the pigment-protein complexes. Moreover, photoprotection requires a chromophore with a triplet energy below that of singlet oxygen (1 eV). Such an organic molecule must have an extensive  $\pi$ -electron system and will therefore strongly absorb visible light.



**Fig. 3** Transient changes in absorbance of II after 353 nm laser flash. The flash was from a Q-switched, frequency tripled neodymium doped laser delivering 30 mJ with a half-peak-height pulse duration of 25 ns. The flash photolysis apparatus has been described previously<sup>26</sup>. Samples were  $\sim 4 \times 10^{-6}$  M in benzene solution and were deaerated by bubbling argon. a, Carotenoid triplet absorption at 535 nm, 1  $\mu\text{s}$  per div., 6.1% change in transmission ( $\delta T$ ) per div. b, Porphyrin triplet absorption at 440 nm, 1  $\mu\text{s}$  per div., 6.8%  $\delta T$  per div.; c, carotenoid triplet absorption at 535 nm, 5  $\mu\text{s}$  per div., 6.1%  $\delta T$  per div.



**Fig. 4** Transient changes in absorbance of I after 353 nm laser flash. All conditions as in Fig. 3 legend. *a*, Carotenoid triplet absorption at 535 nm, 100 ns per div., 2.4%  $\delta T$  per div. (inset is laser pulse swept at 20 ns per div.); *b*, absorption at 440 nm, 100 ns per div., 2.8%  $\delta T$  per div. This change is due to depletion of the carotenoid ground state. *c*, Carotenoid triplet absorption at 535 nm, 1  $\mu$ s per div., 4.4%  $\delta T$  per div.; *d*, absorption at 440 nm, 1  $\mu$ s per div., 2.8%  $\delta T$  per div. This change is due to the repopulation of the carotenoid ground state.

The occurrence of a carotenoid-chlorophyll-protein ultra-structure which collects this light for photosynthesis while providing photoprotection is characteristic of the economy of biological systems.

Finally, biomimetic synthetic solar energy conversion systems using chlorophylls or porphyrins may well inherit problems with light-gathering efficiency and photoprotection similar to those encountered *in vivo*. As demonstrated above, carotenoid polyenes might provide a biomimetic solution to these problems.

This material was prepared with the partial support of the US Department of Energy, grant DE-FG02-79ER10545. This work was also supported by grants from the Cancer Research Campaign and the MRC. We thank Hoffmann-La Roche, Inc., for a gift of  $\beta$ -apo-8'-carotenal.

Received 29 October; accepted 15 December 1980.

- Goodwin, T. W. in *Chemistry and Biochemistry of Plant Pigments* Vol. 1 (ed. Goodwin, T. W.) 228 (Academic, London, 1976).
- Griffiths, M., Siström, W. R., Cohen-Bazire, G. & Stanier, R. Y. *Nature* **176**, 1211-1214 (1955).
- Cohen-Bazire, G. & Stanier, R. Y. *Nature* **181**, 250-252 (1958).
- Krinsky, N. I. in *Photophysiology* Vol. 3 (ed. Giese, A. C.) 123-195 (Academic, New York, 1968).
- Wolff, Ch. & Witt, H. T. Z. *Naturforsch.* **24b**, 1031-1037 (1969).
- Monger, T. G., Cogdell, R. J. & Parson, W. W. *Biochim. biophys. Acta* **449**, 136-153 (1976).
- Kung, M. C. & DeVault, D. *Photochem. Photobiol.* **24**, 87-91 (1976).
- Renger, G. & Wolff, Ch. *Biochim. biophys. Acta* **460**, 47-57 (1977).
- Boucher, F., Van Der Rest, M. & Gingras, G. *Biochim. biophys. Acta* **461**, 339-357 (1977).
- Mathis, P., Butler, W. L. & Satoh, K. *Photochem. Photobiol.* **30**, 603-614 (1979).
- Krinsky, N. I. in *Carotenoids* (eds Isler, O., Guttman, G. & Solms, U.) 669-716 (Birkhäuser, Basel, 1971).
- Foote, C. S., Chang, Y. C. & Denny, R. W. J. *Am. chem. Soc.* **92**, 5216-5217 (1970).
- Koka, P. & Song, P. S. *Biochim. biophys. Acta* **495**, 220-231 (1977).
- Bazzaz, M. B. & Govindjee, P. *Physiol.* **52**, 257-262 (1973).
- Goedheer, J. D. *Biochim. biophys. Acta* **172**, 252-265 (1969).
- Govindjee & Govindjee, R. in *Bioenergetics of Photosynthesis* (ed. Govindjee) 2-50 (Academic, New York, 1975).
- Dirks, G., Moore, A. L., Moore, T. A. & Gust, D. *Photochem. Photobiol.* **32**, 277-280 (1980).
- Moore, A. L., Dirks, G., Gust, D. & Moore, T. A. *Photochem. Photobiol.* **32**, 691-696 (1980).
- Razi Naqvi, K. *Photochem. Photobiol.* **31**, 523-524 (1980).
- Dexter, D. L. J. *Chem. Phys.* **21**, 836-850 (1953).
- Matheson, I. B. C. & Lee, J. *Chem. Phys. Lett.* **7**, 475-476 (1970).
- Bensasson, R. & Land, E. J. in *Photochemical and Photobiological Reviews* (ed. Smith, K.) Vol. 3, 163-191 (Plenum, New York, 1978).
- Bensasson, R., Dawe, E. A., Long, D. A. & Land, E. J. *JCS Faraday Trans. I* **73**, 1319-1325 (1977).
- Bensasson, R., Land, E. J. & Maudinas, B. *Photochem. Photobiol.* **22**, 189-193 (1976).
- Beddard, G. S., Davidson, R. S. & Tretheway, K. R. *Nature* **267**, 373-374 (1977).
- Bensasson, R., Chachaty, C., Land, E. J. & Salet, C. *Photochem. Photobiol.* **16**, 27-37 (1972).
- Dallinger, R. F., Woodruff, W. H. & Rogers, M. A. J. *Photochem. Photobiol.* **33**, 275-277 (1981).

## Absence of ocular dominance patches in dark-reared cats

N. V. Swindale

Physiological Laboratory, Downing Street,  
Cambridge CB2 3EG, UK

If a young monkey or kitten is monocularly deprived for a period of days or weeks, the ocular dominance stripes or patches formed in layer IV of the visual cortex by the geniculo-cortical afferents driven by that eye become smaller, while the patches formed by afferents from the other, experienced eye, spread out and increase in size<sup>1,2</sup>. One explanation for this effect is that it results from a disturbance of a competitive process which, during the first weeks of life, guides a 'sorting out' of the initially intermixed right and left eye inputs into complementary, largely non-overlapping territories<sup>1,3</sup>. One feature of this process may be a local interaction between right and left eye synapses in which like synapses reinforce each other's growth rates and cause rejection of the other eye's synapses<sup>1,4</sup>. If this is the case, then the effect of monocular deprivation on the relative sizes of the two sets of columns can be explained by supposing that the strengths of the effects exerted by the deprived eye are reduced<sup>1,4</sup>. This explanation has a testable consequence: if both eyes are deprived of vision then each eye should be made less effective in eliminating the other eye's inputs, and the overall rate at which the ocular dominance columns form should be decreased<sup>4</sup>. Although LeVay *et al.*<sup>5</sup> found that columns were present in a 7-week old monkey reared in the dark from the age of 3 days, this result does not necessarily imply that the rate of column formation had been normal, because in normal monkeys the columns are well developed by 3 or more weeks of age<sup>5</sup>. I report here the results of transneuronal autoradiography in cats, which show that columns, as revealed anatomically, are undetectable in most parts of the visual cortex of cats reared in the dark for periods of up to 20 weeks, implying that visual experience is necessary for their proper formation.

Six kittens were used; four were reared in complete darkness and studied at the ages of 8, 10, 16 and 20 weeks, while two normal kittens of 8 and 10 weeks of age served as controls. Each animal had its right eye injected, under Halothane anaesthesia, with 2 mCi of <sup>3</sup>H-proline dissolved in 100  $\mu$ l of saline. They survived a further 12-14 days, after which they were given a lethal dose of barbiturate and perfused with 4% paraformaldehyde in 0.1 M phosphate buffer. Frozen sections 30  $\mu$ m thick were cut in a horizontal plane through the visual cortex, mounted on gelatin-coated slides, defatted and dipped in Ilford K2 emulsion diluted 1:3 with 1% glycerol<sup>6</sup>. After exposure periods of 2 weeks to 2 months, the slides were developed in D19 (Kodak), fixed in hypo, and coverslipped without counter-staining.

**Table 1** Degree of spillover in the lateral geniculate nuclei ipsilateral and contralateral to the injected eye

| Age (weeks) | Right (ipsi) | Left (contra) |
|-------------|--------------|---------------|
| 8           | 16%          | 26%           |
| 10          | *            | *             |
| 8           | 9%           | 21%           |
| 10          | 9%           | 17%           |
| 16          | *            | *             |
| 20          | 23%          | 12%           |

The table shows the degree of spillover in the lateral geniculate nuclei ipsilateral and contralateral to the injected eye (which was always on the right side) of the normal and experimental animals. The amount of label present in the lamina (A or A1) not innervated by the injected eye is expressed as a percentage of that in the lamina which was innervated. There is no suggestion that spillover is more severe in the dark-reared animals.

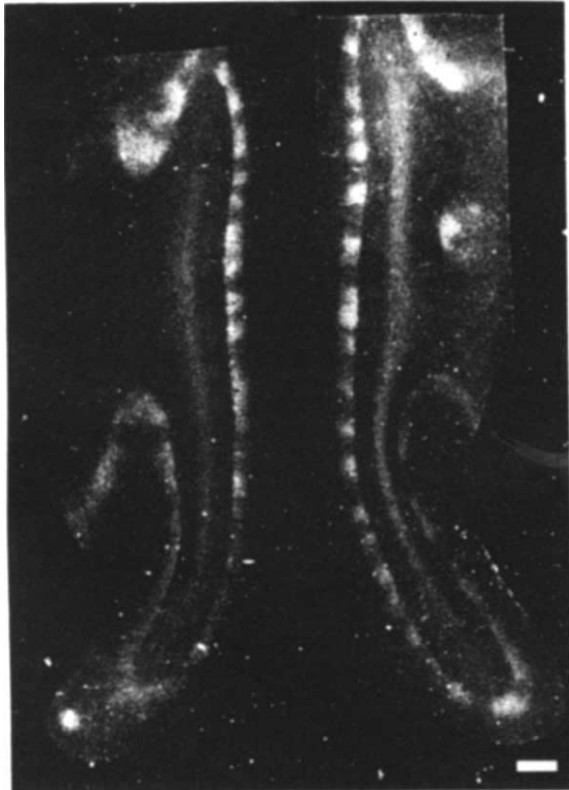
\* Data unavailable.



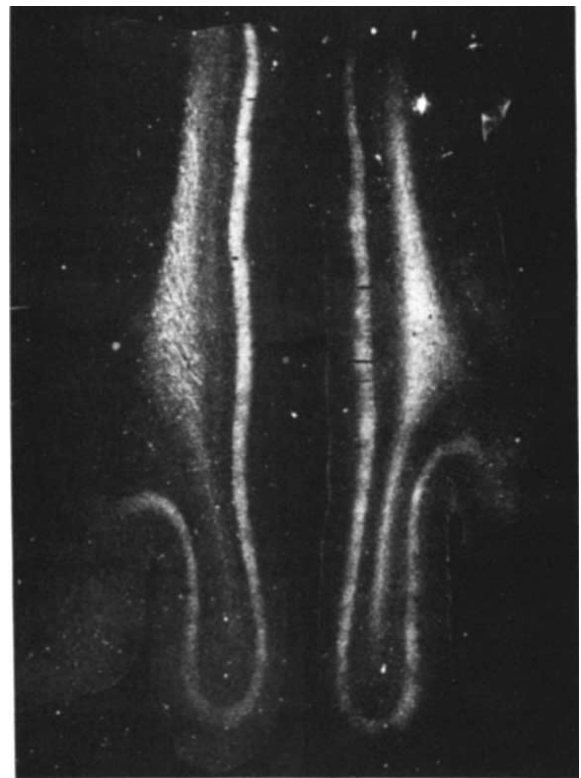
As spillover of radioactive label into geniculate laminae not innervated by the injected eye has been shown to lead to an underestimate of the degree of segregation of the two eyes' terminals in the visual cortex<sup>3</sup>, I also examined the effect of dark rearing on this. The procedure was similar to that of LeVay *et al.*<sup>3</sup>: 1- $\mu$ m Araldite sections of the lateral geniculate nucleus were cut, coated with emulsion and processed in the same way as the frozen sections. Grain counts were made in laminae A and A1, which are innervated by the contralateral and ipsilateral eyes respectively, and the ratio of the two values expressed as a percentage (Table 1). No systematic differences between dark-reared and normal animals were found, although spillover was usually more severe on the contralateral side<sup>3</sup>.

The labelling pattern in both the 8- and 10-week old normal animals studied was similar to that found in normal adult animals<sup>3,7</sup> (Fig. 1). In layer IV of area 17, ipsilateral to the injected eye, the label formed patches 200–600  $\mu$ m wide, spaced at roughly 800- $\mu$ m intervals, with a level of label in the gaps between which was only slightly higher than the background in adjacent layers of the cortex. In area 17 contralateral to the injected eye, the label was also distributed periodically, although the level of label in the gaps between the columns was higher than on the ipsilateral side. This probably results from the greater degree of spillover on this side.

All four dark-reared animals gave similar results. Over most of area 17, on sides both ipsilateral and contralateral to the injected eye, the distribution of label was almost uniform (Fig. 2), although in some parts of the cortex, slight fluctuations, whose periodicity was not obviously abnormal, were present. The level of label in the troughs of these fluctuations was always much higher than in the gaps between the labelled columns in the corresponding hemispheres of normal animals. They were always more pronounced in the ipsilateral hemisphere and were less noticeable or absent in regions of cortex receiving inputs



**Fig. 1** Dark field autoradiographs from single horizontal sections through the left and right visual cortices of an 8-week old normal kitten whose right eye was injected with radioactive label. The label has a periodic, patchy distribution in layer IV of both hemispheres. In this and Fig. 2, the left hemisphere is on the left-hand side, and anterior is upmost. Scale bar, 1 mm.



**Fig. 2** Autoradiographs from the left and right hemispheres of an 8-week old dark-reared kitten, whose right eye was injected. The distribution of label is nearly uniform in both hemispheres.

from peripheral retina. Contrary to expectation they were, if anything, less noticeable in the older dark-reared animals.

Although these results show that total visual deprivation interferes substantially with the process by which left and right eye terminals segregate to form ocular dominance columns, it remains unclear whether the deprivation is, as predicted, simply slowing the rate of segregation, in which case the columns should form eventually, or whether the process halts or even reverses after an intermediate stage of development. Some degree of development has occurred in the dark, because the fluctuations found in some parts of the cortex are absent from normal 2-week old animals<sup>3</sup> (despite their several days' visual experience). The dissimilarity between these results from cats and those from the 7-week old dark-reared monkey studied by LeVay *et al.*<sup>5</sup>, which possessed an apparently normal set of ocular dominance columns, is also puzzling, but more experimental work is needed to clarify the nature of the difference.

Cats reared in the dark for four or more months are known to be abnormally sensitive to the effects of monocular deprivation, judged by changes in the responsiveness of cortical cells to visual stimuli<sup>8</sup>. This is at a time when similar deprivation would be without effect in a normal animal. The present results therefore support the idea that the critical period for monocular deprivation is linked to the time when left and right eye inputs to the cortex are incompletely segregated<sup>1,3,4</sup>.

I thank Lyn Cummings for technical assistance and Ralph Freeman for helpful discussions. The research was supported by grants to Professor H. B. Barlow from the MRC and the Royal Society.

Received 10 December 1980; accepted 4 February 1981.

- Hubel, D. H., Wiesel, T. N. & LeVay, S. *Phil. Trans. R. Soc. B* **278**, 131–163 (1977).
- Shatz, C. J. & Stryker, M. P. *J. Physiol., Lond.* **281**, 267–283 (1978).
- LeVay, S., Stryker, M. P. & Shatz, C. J. *J. comp. Neurol.* **179**, 223–244 (1978).
- Swindale, N. V. *Proc. R. Soc. B* **208**, 243–264 (1980).
- LeVay, S., Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **191**, 1–51 (1980).
- Rogers, A. W. *Techniques of Autoradiography* 2nd edn (Elsevier, Amsterdam, 1973).
- Shatz, C. J., Lindström, S. & Wiesel, T. N. *Brain Res.* **131**, 103–116 (1977).
- Cynader, M. in *Developmental Neurobiology of Vision* Series A: Life Sciences Vol. 27 (ed. Freeman, R. D.), 109–120 (Plenum, New York, 1979).

## Common precursor of lysozymes of hen egg-white and bacteriophage T4

B. W. Matthews, M. G. Grütter, W. F. Anderson\* & S. J. Remington\*

Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

**The lysozymes of hen egg-white and bacteriophage T4 have similar catalytic properties but their amino acid sequences are not homologous. The question therefore arises whether they are derived from a common ancestral protein or have arisen independently. On the basis of the data we have gathered, it is shown here that the two enzymes are similar in the conformation of their backbones, in their modes of binding substrates, in specific protein-substrate interactions and in their presumed modes of action. We conclude that the two enzymes have diverged from a common precursor. This seems to be the most convincing example to date of the divergence of proteins with nonhomologous amino acid sequences.**

If anything, the large number of protein structures and amino acid sequences now known has confused rather than simplified the main issue of whether extant proteins have evolved from a small number of ancestral precursors or have arisen independently. When the amino acid sequences are closely homologous, there is no difficulty: three-dimensional structures are always similar, and it is reasonable to infer that there is a common ancestral protein. In the absence of sequence homology, on the other hand, the origin of two proteins is logically ambiguous. Functional and/or structural similarity may indicate either divergent evolution from a common precursor or convergent evolution from amino acid sequences arising independently.

In many circumstances, tell-tale sequence homology is not readily apparent. The most notable example is the 'nucleotide-binding domain' of four NAD-dependent dehydrogenases and certain other enzymes<sup>1-5</sup>. Although the amino acid sequences of lactate, alcohol and glyceraldehyde-3-phosphate dehydrogenases are not obviously homologous, when the sequences are aligned in such a way that common structural features coincide, a degree of sequence homology becomes evident, suggesting divergent evolution. In other cases of enzymes with similar structures but no apparent sequence homology, the question of divergent or convergent evolution has remained open<sup>6-21</sup>. The greater the structural correspondence between two proteins, as estimated by a variety of methods<sup>4,16,22-28</sup>, the more likely are the two proteins to have evolved from the same precursor, but structural similarity *per se* is no proof of divergence. In some instances, two domains of a single protein have been observed to have structural similarity but not sequence homology. The fact that the polypeptide chains of the two domains are contiguous, even though their sequences are not related, has been taken as evidence for gene duplication, that is, a special case of divergent evolution<sup>9,28-30</sup>.

In the case of bacteriophage T4 lysozyme (T4L) and hen egg-white lysozyme (HEWL), there are additional critical features common to the respective structures which strongly suggest that the two enzymes have diverged from a common precursor. Correspondence between HEWL and T4L can be seen at three levels: (1) the respective backbones, (2) the locations of saccharides bound in the respective active site clefts and (3) specific protein-substrate interactions and the probable mechanisms of catalysis.

Rossmann and Argos<sup>4</sup> showed that, allowing for certain insertions and deletions, the amino-terminal half of T4L is structurally similar to HEWL. Of the 129 residues in HEWL and 164 in T4L, 78  $\alpha$ -carbon atoms were found to be 'equivalent' and could be superimposed within a r.m.s. distance of 4.1 Å. Notwithstanding the structural correspondence, the minimum base change per codon for the 78 'equivalent' residues is 1.53, a value expected for two amino acid sequences chosen at random. Remington and Matthews<sup>25</sup> used a different method of comparing the two lysozyme structures and concluded that the agreement between them was significantly better than one would expect by chance. The part of T4L 'equivalent' to HEWL is shown in Fig. 1.

Recently, we have completed a study of the binding of a series of mono- and oligosaccharides to phage lysozyme which has permitted a detailed comparison of the respective active sites of T4L and HEWL<sup>31,32</sup>. We took the coordinates of the tetrasaccharide  $\delta$ -lactone bound to HEWL<sup>31</sup>, and, using the respective transformations between HEWL and T4L proposed by Rossmann and Argos<sup>4</sup> and Remington and Matthews<sup>25</sup>, transformed these coordinates into the T4L coordinate frame. As summarized in Table 1, the centres of the four sugars agree within 1–2 Å with the saccharide binding sites determined experimentally for T4L. The experimental data provide striking confirmation of the sugar binding sites in T4L predicted by Rossmann and Argos, and support their conclusions concerning the relationship between the two lysozymes. Furthermore, there is remarkable agreement (Fig. 2) between parts of T4L and HEWL which are critically important for substrate binding and catalysis.

In the case of HEWL, Phillips and co-workers proposed a mechanism of action in which substrate strain in subsite D favours the formation of a carbonium ion intermediate which is stabilized by Asp 52, with Glu 35 acting as the proton donor<sup>33,34</sup>. Subsequent work suggests that electronic stabilization rather than strain may be more important in promoting catalysis<sup>35,36</sup>. In the case of T4L, we see a close parallel<sup>31,32</sup>. Here, model building suggests that, as with HEWL, a saccharide with the normal chair conformation cannot be accommodated in site D because of steric interference, but a saccharide in the sofa conformation, as in the proposed transition state, will fit into the D site. For HEWL, the critical close contact in site D is between C<sub>(6)</sub> and the backbone C=O of Gln 57. As can be seen in Fig. 2, an exactly analogous close approach occurs for T4L between C<sub>(6)</sub> and the backbone C=O of Gly 30. In addition, for both enzymes the NH of the adjacent peptide forms a characteristic hydrogen bond to the acetamido group of saccharide C which is thought to help promote the preferential binding of monosaccharides in site C.

Finally, Asp 20 of T4L is in virtually the identical position, relative to substrate, as Asp 52, and is ideally placed to stabilize the developing positive charge on the postulated carbonium ion intermediate. Glu 11 of T4L is close to, but not in exactly the same relative position, as Glu 35 of HEWL, and is salt linked to Arg 45, in contrast to the nonpolar environment of Glu 35. Assuming a slight conformational adjustment when substrate is bound, Glu 11 could move into the required position to act as

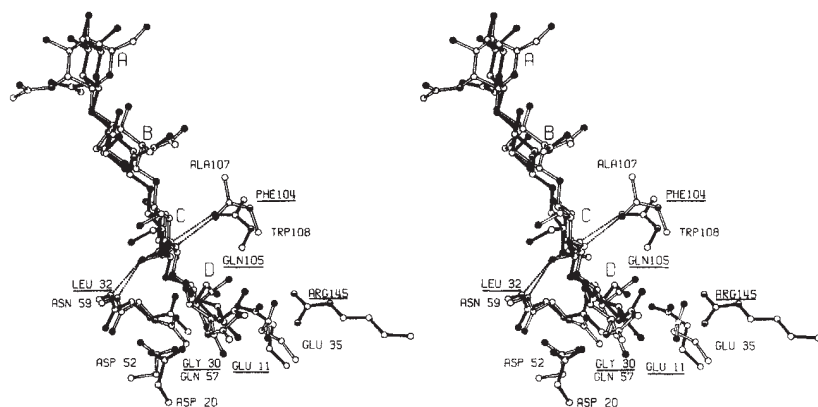
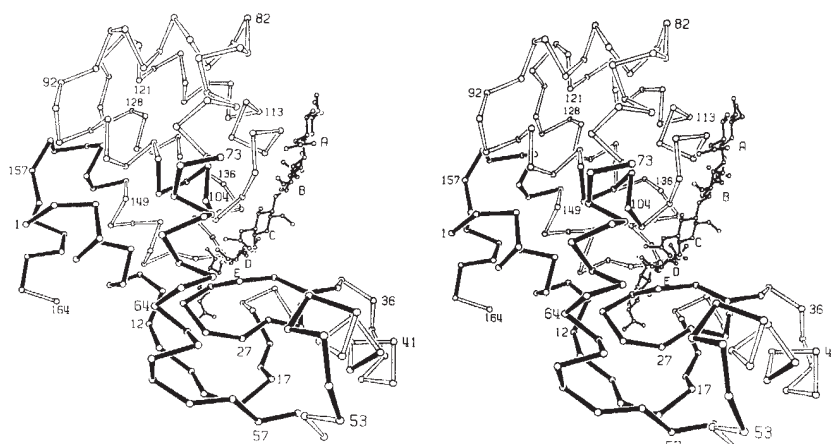
**Table 1** Comparison of predicted and observed saccharide binding sites for phage lysozyme

| Saccharide subsite | Distance between predicted and observed site (Å)                           |                                                                                                      |
|--------------------|----------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
|                    | Prediction based on 78 'equivalent' $\alpha$ -carbon atoms <sup>4,32</sup> | Prediction based on the best agreement between 80 contiguous $\alpha$ -carbon atoms <sup>25,32</sup> |
| A                  | 1.8                                                                        | 2.6                                                                                                  |
| B                  | 1.5                                                                        | 1.6                                                                                                  |
| C                  | 1.5                                                                        | 1.1                                                                                                  |
| D                  | 1.5                                                                        | 1.2                                                                                                  |

\* Present addresses: Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7 (W.F.A.); Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG (S.J.R.).



**Fig. 1** Stereo view of the  $\alpha$ -carbon backbone of bacteriophage T4 lysozyme showing the location of a bound oligosaccharide. The solid lines connect the 78  $\alpha$ -carbon atoms of phage lysozyme which have equivalent  $\alpha$ -carbons in hen egg-white lysozyme<sup>4</sup>.



**Fig. 2** Superposition of elements of the active site of hen egg-white lysozyme (open bonds), including a bound tetrasaccharide, on the corresponding elements of T4 phage lysozyme (solid bonds and names underlined).

the proton donor, as Glu 35 is postulated to do in HEWL. (Additional details are given in ref. 32.)

Thus, phage lysozyme and HEWL have similarities in their backbone conformations, in the locations of bound substrates and in the specific details of substrate binding and catalysis. These similarities strongly suggest that the two lysozymes have diverged from a common precursor. To argue for convergent, rather than divergent evolution, one would have to postulate that the similarities between the two lysozymes which we have enumerated all arose independently. We consider this unlikely.

The two lysozymes provide an example of the changes in proteins which can occur during evolution. In this case the only amino acids which seem to have retained their identities are Glu 11 (Glu 35) and Asp 20 (Asp 52). Other similarities between the respective lysozymes (Figs 1, 2) are at the level of the protein backbone rather than side-chain identity. As will be described in more detail elsewhere, the part of the upper lobe of T4L (Fig. 1), which has no counterpart in HEWL, seems to be involved in binding the cross-linking peptide of *Escherichia coli* cell walls. This peptide is required for substrates of T4L, but is not essential for HEWL.

After the submission of this manuscript, Jung *et al.*<sup>37</sup> reported the locations of the four exons of HEWL. Exon 2 codes for amino acids 28–32 and exon 3 for amino acids 82–108. Taken together, these two regions include the part of HEWL which is most similar to T4L and support the hypothesis<sup>38</sup> that exons may have been involved in the evolution of proteins. The same suggestion has been made independently by P. J. Artymuk, C. C. F. Blake and A. E. Sippel (personal communication).

This work was supported in part by a Swiss NSF Fellowship to M.G.G., a NIH Postdoctoral Fellowship (GM 05972) to W.F.A., NIH research grants GM 20066 and GM 21967, NSF grant PCM-8014311 and a grant from the M. J. Murdock Charitable Trust. We thank Dr Colin Blake for communicating his manuscript before publication.

Received 15 September 1980; accepted 5 January 1981.

- Blake, C. C. F. *Nature* **250**, 284–285 (1974).
- Blake, C. C. F. *Nature* **267**, 482–483 (1977).
- Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250**, 194–199 (1974).
- Rossmann, M. G. & Argos, P. *J. molec. Biol.* **105**, 75–96 (1976).
- Rossmann, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. in *The Enzymes* Vol. 11, 3rd edn (ed. Boyer, P. E.) 61–102 (Academic, New York, 1975).
- Blake, C. C. F. & Evans, P. R. *J. molec. Biol.* **84**, 585–601 (1974).
- Banks, R. D. *et al. Nature* **279**, 773–777 (1979).
- Campbell, J. W., Watson, H. C. & Hodgson, G. I. *Nature* **250**, 301–303 (1974).
- Ploegman, J. H., Drent, G., Kalk, K. H. & Hol, W. G. J. *J. molec. Biol.* **123**, 557–594 (1978).
- Fletterick, R. J., Syguch, J., Semple, H. & Madsen, N. B. *J. biol. Chem.* **251**, 6142–6146 (1976).
- Weber, I. T. *et al. Nature* **274**, 433–437 (1978).
- Levine, M., Muirhead, H., Stammers, D. K. & Stuart, D. I. *Nature* **271**, 626–630 (1978).
- Eichele, G. *et al. J. molec. Biol.* **133**, 161–180 (1979).
- Borisov, V. V., Borisova, S. N., Sosfenov, N. I. & Vainshtein, B. K. *Nature* **284**, 189–190 (1980).
- Matthews, D. A. *et al. J. biol. Chem.* **253**, 6946–6954 (1978).
- Rossmann, M. G. & Argos, P. *J. molec. Biol.* **109**, 99–129 (1977).
- Rossmann, M. G. & Argos, P. *J. biol. Chem.* **250**, 7525–7532 (1975).
- Richardson, J. S., Richardson, D. C., Thomas, K. A., Silvertown, E. W. & Davies, D. R. *J. molec. Biol.* **102**, 221–235 (1976).
- Phillips, D. C., Sternberg, M. J. E., Thornton, J. M. & Wilson, I. A. *J. molec. Biol.* **119**, 329–351 (1978).
- Argos, P., Rossmann, M. G. & Johnson, J. E. *Biochem. biophys. Res. Commun.* **75**, 83–86 (1977).
- Blow, D. M., Irwin, M. J. & Nyborg, J. *Biochem. biophys. Res. Commun.* **76**, 728–734 (1977).
- Schulz, G. E. & Schirmer, R. H. *Nature* **250**, 142–144 (1974).
- Richardson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2619–2623 (1976).
- Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **105**, 357–382 (1976).
- Remington, S. J. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2180–2184 (1978).
- Remington, S. J. & Matthews, B. W. *J. molec. Biol.* **140**, 77–99 (1980).
- Schulz, G. E. *J. molec. Evol.* **9**, 339–342 (1977).
- Schulz, G. E. *J. molec. Biol.* **138**, 335–347 (1980).
- McLachlan, A. D. *J. molec. Biol.* **128**, 49–79 (1979).
- McLachlan, A. D. in *Protein Folding* (ed. Jaenicke, R.) 79–99 (Elsevier, Amsterdam, 1980).
- Anderson, W. F., Grütter, M. G., Remington, S. J. & Matthews, B. W. *J. molec. Biol.* (in the press).
- Matthews, B. W., Remington, S. J., Grütter, M. G. & Anderson, W. F. *J. molec. Biol.* (in the press).
- Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tjian, R. *J. molec. Biol.* **88**, 349–371 (1974).
- Blake, C. C. F. *et al. Proc. R. Soc. B* **167**, 378–388 (1967).
- Warshel, A., Levitt, M. *J. molec. Biol.* **103**, 227–249 (1976).
- Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5250–5254 (1978).
- Jung, A., Sippel, A. E., Grez, M. & Schutz, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5759–5763 (1980).
- Gilbert, W. *Nature* **271**, 501 (1978).

## Unusual filterable oncogenic agent isolated from horizontally transmitted Syrian hamster lymphomas

Joseph H. Coggin Jr\*, John E. Oakes\*,  
Robert J. Huebner† & Ray Gilden‡

\* Department of Microbiology and Immunology, University of South Alabama, College of Medicine, Mobile, Alabama 36688, USA

† National Cancer Institute, Building 37, Room 2D28, Bethesda, Maryland 20205, USA

‡ Frederick Cancer Center, Frederick, Maryland 21701, USA

**An outbreak of horizontally transmitted malignant lymphoma in an experimental hamster holding facility was previously reported<sup>1</sup>. Retroviridae (oncornavirus) or other conventional oncogenic viruses (oncoviruses) could not be detected in these lymphomas by immunological methods, direct isolation procedures or electron microscopy<sup>2</sup> but an infectious agent was clearly involved. The incidence of lymphomas during five recurrent epidemics ranged from 50 to >90% in young, inbred and random-bred Syrian golden hamsters exposed<sup>1</sup>. The agent seemed to be resistant to UV inactivation, formaldehyde vapour and other viricidal agents (chlorine and iodine), and stable for long periods in the absence of hamster hosts in the contaminated facility<sup>1,2</sup>. Associated disease syndromes in exposed hamsters included severe enteritis, pyelonephritis, the occasional appearance of warts, poor breeding efficiency and intussusception<sup>1,2</sup>. We now report the successful, cell-free isolation of an unusual, filterable agent prepared in protamine sulphate buffer from primary and animal-passaged lymphomas, which produces lymphomas with good efficiency when injected subcutaneously (s.c.) into newborn Syrian inbred (LSH) and random-bred (LVG) hamsters. The agent could be reisolated from these induced lymphomas and injected into other hamsters to reproduce the neoplastic condition. It showed characteristics suggested for a mammalian viroid (a non-encapsidated, DNase-sensitive low-molecular-weight, disease-causing, self-replicating, naturally infectious nucleic acid).**

Various isolation techniques<sup>3-6</sup> were used in attempts to obtain cell-free, filterable agent responsible for lymphoma induction in young or newborn hamsters inoculated either s.c. or intraperitoneally (i.p.) (Table 1). None of these yielded cell-free oncogenic agent, but the incidence of 'background' lymphoma mediated by naturally transmitting agent among the hamsters housed in the contaminated facility was suppressed in hamsters challenged with non-oncogenic extracts of lymphoma, compared with placebo controls which received buffer or normal spleen extracts. In this report, data are given for the incidence of background lymphomas detected among control or test hamsters placed in the agent-contaminated facility. The incidence of background lymphomas can range from 0 to >50% in some rooms. Such spontaneous lymphomas contracted from agent in the facility were unavoidable because we could not use new facilities for each experiment. Previous work demonstrated that it was difficult to rid the unit of contaminating agent by extensive sanitization that would destroy most viruses. However, background lymphomas were readily discerned from protamine sulphate (PS)-stabilized agent induced tumours by site of appearance.

We obtained cell-free agent using a simple extraction method in which PS (0.01 mg ml<sup>-1</sup>) was added to the extraction buffer (Fig. 1). Filtered (0.45–0.1 µm) extracts of lymphoma prepared with PS were able to produce 40–95% lymphomas in trials carried out over 2 yr when inoculated s.c. into newborn hamsters. Two additional PS preparations were tested s.c. in newborns in a new facility which was free of contamination and

**Table 1** Protection of hamsters against contagious lymphoma after injection of Moloney extracts of spontaneous intestinal lymphoma

| Preparation         | Hamsters with lymphomas/survivors<br>Test extract (%) | Control buffer<br>solution (%) |
|---------------------|-------------------------------------------------------|--------------------------------|
| Moloney extract:    |                                                       |                                |
| Trial 1             | 4/18 (22)                                             | 9/9 (100)                      |
| Trial 2             | 0/16 (0)                                              | 4/11 (36)                      |
| Trial 3             | 2/34 (6)                                              | 3/9 (33)                       |
| Freeze-thaw extract | 3/34 (9)                                              | 4/9 (44)                       |
| Saline homogenate   | 1/11 (10)                                             | 3/9 (33)                       |
| Phenol extraction*  |                                                       |                                |
| Trial 1             | 0/20                                                  | 0/4                            |
| Trial 2             | 0/30                                                  | 0/7                            |
| Trial 3             | 0/37                                                  | NT                             |
| Hirt extraction*    |                                                       |                                |
| Trial 1             | 0/10                                                  | 2/10 (20)                      |
| Trial 2             | 0/23                                                  | 1/9 (11)                       |

The Moloney extract comprised 15 g of lymphoma extracted in citrate buffer and hyaluronidase (for methods see Fig. 1 legend). 0.5 ml was injected into the right rear legs of 3–4-week-old LVG hamsters. All solutions were passed through a 0.45-µm filter. All lymphomas recorded were at node sites in the peritoneal cavity or in the cervical-thymic node area. The freeze-thaw extract comprised 15 g of lymphoma extracted after three repetitive flash freezes (–80 °C) and slow thaw (25 °C) in phosphate-buffered saline (PBS), pH 6.8. The solution was passed through a 0.45-µm filter and 0.05 ml injected i.p. The saline homogenate comprised 15 g of lymphoma extracted by homogenization in PBS and filtered through a 0.45-µm membrane. 0.05 ml was injected i.p. Each group originally comprised 40–48 newborns. The survivors were rapidly reduced to those given by enteritis. Young hamsters which died from enteritis had no lymphomas at the time of death (2–6 weeks). Enteritis was not seen after 5–6 weeks, and was observed equally in control groups and in hamsters injected with lymphoma extracts. These trials were carried out in a new trailer unit. NT, not tested.

\*Lymphomas in these groups were detected in the lower flank, adjacent to the injection site.

were oncogenic. These tumour preparations were from cell-free, PS agent induced lymphomas. When the same PS-stabilized extracts were injected i.p., there were fewer tumours in the injection area than after s.c. injection (data not shown).

Electron microscopy using several staining methods<sup>7,8</sup> on spontaneous lymphomas, animal-passaged lymphomas and concentrated PS extracts of lymphoma oncogenic in newborn hamsters were devoid of virions unique to the lymphomas. R or H particles and rare budding particles were infrequently observed in normal and fibroblastic tissue and epithelial cells from lymphomatous hamsters, but these were also commonly detected in non-exposed hamsters<sup>9</sup> from breeder sources. We concluded that there were no virions or virus-like particles (VLPs) unique to lymphoma tissue or to uninvolved hair follicles, skin, spleen, gut, or normal lymphocytes of lymphoma-bearing hamsters.

Preparations of lymphoma or serum from lymphoma-bearing hamsters were tested for mycoplasma, antibody to lymphocytic choriomeningitis virus, SV40, polyoma and SV5 and Reo virus, with negative results. Antibody to Sendai virus was present in healthy, breeder-supplied hamsters. In repeated trials, serum from either lymphoma-bearing or their exposed, disease-free cage-mates failed consistently to neutralize PS-stabilized agent and prevent lymphoma induction.

Hamster age was most important in maximum sensitivity to s.c. oncogenic challenge with PS-stabilized, filtered extracts of lymphoma (Table 2), with 1-day-old, but not 10-day-old, hamsters being most susceptible. The oncogenic agent in PS was readily inactivated by DNase but not by RNase or proteases (Table 3).

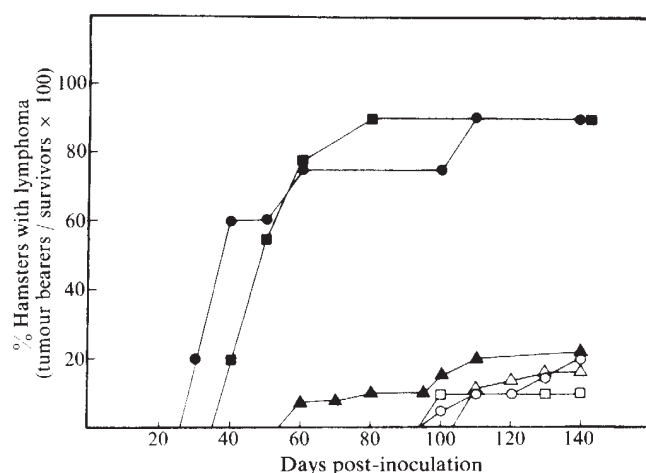
Diener<sup>10</sup> concluded that a viroid etiology should be considered when: (1) no microorganism was consistently associated with the disease; (2) VLPs could not be identified in extracts or sections of diseased tissue by electron microscopy; (3) the



infectious agent was difficult to pellet by ultracentrifugation; and (4) nuclease treatment inactivated the agent. Fulfilment of these criteria alone does not constitute absolute proof of a viroid aetiology which depends on demonstrating non-encapsidated nucleic acid<sup>10</sup>.

We suggest that this lymphoma agent is a naturally transmitting, oncogenic, naked DNA viroid-like agent whose origin and mode of transmission remain unclear. The agent is difficult to sediment in buffer, DNase labile, produces no unique virus particles in lymphoma tissue or normal tissues of the tumour-bearing host nor in concentrated, filtered oncogenic extracts, and requires PS for cell-free stabilization. It seems to be unrelated to known oncogenic Retroviridae or papovaviruses, and generates no neutralizing antibody in trials using tumour-bearer or disease-free cage-mate serum. We must still prove the existence of infectious nucleic acid to claim a viroid status for the oncogenic agent.

Graffi *et al.*<sup>11-13</sup> reported a similar disease in Syrian hamsters and attributed the cause to an unidentified budding virus of the Retroviridae family. They suggested that various agents, such as



**Fig. 1** Lymphoma production using extracts of hamster lymphoma. ○, Control (no inoculations). Hamsters inoculated s.c. in rear flank: ▲, Moloney method; ■, PS method (LSH recipients); ●, PS method (LVG recipients); △, control (Moloney extraction buffer only); □, control (PS solution only). Modified Moloney extraction method: a 6.6% suspension was prepared of intestinal lymphoma which had been finely minced in 0.153 M potassium citrate containing 1 mg of hyaluronidase per 100 ml. This was digested for 1 h at 25 °C with occasional mixing in a blender (15-s pulse) and homogenized in the blender in an ice bath (3 min). The extract was centrifuged at 2,500 r.p.m. for 20 min at 4 °C and the middle supernatant ( $S_1$ ) collected. This was centrifuged again as above and the middle supernatant ( $S_2$ ) collected.  $S_2$  was centrifuged at 17,000g for 1 min at 4 °C and the supernatant ( $S_3$ ) collected.  $S_3$  was filtered through a 8- $\mu$ m Biorad filter, then through a 1- $\mu$ m Biorad filter, and finally a 0.45- $\mu$ m Millipore filter. The  $S_3$  filtrate was centrifuged at 40,000g for 1 h at 4 °C and the pellet rinsed three times with 0.05 M sodium citrate buffer, pH 6.7. The pellet was resuspended in citrate buffer using 1 ml per g of original lymphoma and exposed to three 1-s bursts with a sonicator to achieve cavitation. 0.1 ml was injected s.c. Protamine sulphate method: lymphoma (10–15 g) was minced to a fine suspension in Hank's balanced salt solution, and the final suspension forced through an 18-gauge needle. The cells were pelleted at 1,500 r.p.m. for 10 min at 4 °C and the pellet washed twice with 5 ml PBS, pH 6.8, per 5 g tumour used. The pellet was resuspended in PBS containing 0.01 mg ml<sup>-1</sup> PS, 0.01 M Tris, 0.001 M EDTA (pH 8.0) and kept at 4 °C for 30 min. This was homogenized in a Dounce unit and centrifuged at 2,000 r.p.m. for 10 min at 4 °C. The supernatant ( $S_1$ ) was centrifuged in a 50 Ti Beckman rotor at 10,000g ( $\omega^2 t = 365$ ) for 60 min at 4 °C. The supernatant ( $S_2$ ) was filtered through a 0.45- $\mu$ m Millipore filter, then a 0.22- $\mu$ m filter. 0.1 ml was injected s.c.

**Table 2** Relationship of age to susceptibility to lymphoma induction by cell-free, filtered preparations from spontaneous lymphoma

| Protamine extract of: | Age of challenge (days post-birth) | Final tumour incidence (tumours/survivors, %) |
|-----------------------|------------------------------------|-----------------------------------------------|
| Normal hamster spleen | 1                                  | 2/17 (12)                                     |
|                       | 5                                  | 1/12 (8)                                      |
|                       | 10                                 | 3/22 (9)                                      |
| Buffer alone          | 1                                  | 3/23 (13)                                     |
|                       | 5                                  | 0/11 (0)                                      |
|                       | 10                                 | 2/20 (10)                                     |
| Spontaneous lymphoma  | 1                                  | 14/20 (70) $P < 0.01^*$                       |
|                       | 5                                  | 9/22 (41) $P < 0.05$                          |
|                       | 10                                 | 5/24 (21) $P > 0.05$ NS                       |

A spontaneous lymphoma which appeared in the lymph node of the neck area (weight 5.3 g) was removed and extracted using the PS extraction method (see Fig. 1 legend). 0.05 ml of the final supernatant was injected s.c. into the lower rear quadrant of hamsters. Hamsters were challenged with extracts on the day indicated. All mothers were obtained from the Charles River Hamstery as primiparous and all groups were weaned and sex segregated at 21 days of age. Hamsters were palpated and examined weekly for evidence of tumour. The experiments were terminated at 222 days after birth and survivors autopsied for evidence of tumours. The last tumour was detected at 178 days and the first appeared at 28 days. NS, not significant.

\* Significance by  $\chi^2$  compared with normal spleen cell extract.

**Table 3** Susceptibility of oncogenic agent-producing lymphomas to DNase

| Treatment                                                                   | Cumulative no. lymphomas/<br>no. survivors<br>Days post-injection of neonate<br>LVG hamsters |      |       |      |
|-----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|------|-------|------|
|                                                                             | 0                                                                                            | 80   | 100   | 150  |
| None—filtered PS-stabilized lymphoma extract alone (30 min, 37 °C)          | 0/17                                                                                         | 4/9  | 7/9   | 8/9  |
| DNase (30 min, 37 °C) plus filtered PS-stabilized lymphoma extract          | 0/20                                                                                         | 0/16 | 0/16  | 0/15 |
| Trypsin (0.25%; 20 min, 37 °C) plus filtered PS-stabilized lymphoma extract | 0/16                                                                                         | 8/16 | 16/16 | —    |
| RNase (30 min, 37 °C) plus filtered PS-stabilized lymphoma extract          | 0/10                                                                                         | 3/9  | 4/6   | 6/6  |
| Heat-inactivated DNase* (30 min, 37 °C) plus filtered PS-stabilized extract | 0/13                                                                                         | 3/12 | 8/9   | 9/9  |
| Control—PS-containing extraction solution only (no tumour)                  | 0/15                                                                                         | 0/14 | 1/14  | 1/14 |

DNase, deoxyribonuclease 1, type 1, bovine pancreas (Sigma); RNase free, 250  $\mu$ g ml<sup>-1</sup>. As little as 50  $\mu$ g ml<sup>-1</sup> gave the same result in another trial. Trypsin, type IX (Sigma), 0.5%. Papain was also ineffective in another trial. RNase, ribonuclease, type A-1, protease free (Sigma).

\* DNase was inactivated by boiling for 10 min before use.

human tumour DNA and hamster papilloma DNA, could activate an endogenous C-type virus which induced the lymphomas observed. We find no evidence for such a retrovirus in our epidemic disease. M. Gardner (personal communication) found that saline could activate lymphoma in Graffi's hamster stock; therefore it is possible that an agent identical to the one described here was being transmitted unrecognized in the Graffi hamster.

There are several problems in resolving the nucleic acid nature of this lymphoma-inducing agent: (1) it is highly stable in the environment (facility contamination); (2) the bioassay used

to detect the agent is long-term and expensive; (3) agent-producing cell cultures of the lymphoma have not been obtained despite extensive culture efforts—cells fail to grow *in vitro*; (4) reliable *in vitro* transformation of target cells by the agent has not been detected; (5) antigenic markers for the agent on transformed lymphoma cells are not yet sufficiently developed; and (6) other potential mammalian hosts are not susceptible to the agent. Work is being done to characterize this infectious, oncogenic DNA agent.

We thank Dr T. O. Diener for helpful comments during preparation of the manuscript. This work was supported by NCI grant CA 23491, NIH.

Received 5 September 1980; accepted 2 January 1981.

1. Ambrose, K. R. & Coggin, J. H. *J. natn. Cancer Inst.* **54**, 877–880 (1975).
2. Coggin, J. H., Thomas, K. V. & Huebner, R. *Fedn Proc.* **70**, 2086–2088 (1978).
3. Moloney, J. B. *J. natn. Cancer Inst.* **24**, 933–937 (1960).
4. Moloney, J. B. *Fedn Proc.* **21**, 19–21 (1962).
5. Green, M., Wold, W. S. M., Mackey, J. K. & Rigden, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6606–6610 (1979).
6. Hirt, B. *J. molec. Biol.* **26**, 365–369 (1967).
7. Fresca, J. M. & Parks, V. R. *J. Cell Biol.* **25**, 157–163 (1965).
8. Huxley, H. E. & Zubay, G. *J. biophys. biochem. Cytol.* **11**, 273–278 (1961).
9. Bernhard, W. & Tournier, P. *Annls Inst. Pasteur Paris* **107**, 447–452 (1964).
10. Diener, T. O. *Viroids and Viroid Diseases* 1–224 (Wiley, New York, 1979).
11. Graffi, A., Bender, E., Schramm, T., Kuhn, W. & Schneiders, F. *Proc. natn. Acad. Sci. U.S.A.* **64**, 1172–1175 (1969).
12. Graffi, A. *et al. Br. J. Cancer* **22**, 577–581 (1968).
13. Graffi, A., Schramm, T., Graffi, I., Bierwolf, D. & Bender, E. *J. natn. Cancer Inst.* **40**, 867–873 (1968).

## $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ has one functioning phosphorylation site per $\alpha$ subunit

W. H. M. Peters, H. G. P. Swarts,  
J. J. H. M. de Pont,  
F. M. A. H. Schuurmans Stekhoven  
& S. L. Bonting

Department of Biochemistry, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

$(\text{Na}^+ + \text{K}^+)\text{ATPase}$  contains two different subunits, a catalytic subunit ( $\alpha$ ) and a subunit with uncertain function ( $\beta$ ). The enzyme binds ATP, ouabain and vanadate, and can be phosphorylated by ATP as well as by inorganic phosphate. From the previously reported maximal binding and phosphorylation capacities of 3.5–4.3 nmol P per mg protein<sup>1–3</sup> (based on Lowry protein determination) and the earlier molecular weight value of ~250,000<sup>1,2</sup>, a molar binding and phosphorylation capacity of 0.87–1.07 mol per mol enzyme was derived. As it is generally agreed that the enzyme molecule contains two  $\alpha$  subunits<sup>1–6</sup> or even a multiple of two<sup>7,8</sup>, it has been suggested that the enzyme operates by means of a so-called 'half-of-the-sites mechanism', whereby only one of the two  $\alpha$  subunits can be phosphorylated at any one time<sup>9,10</sup>. We now present evidence that every  $\alpha$  subunit can be phosphorylated simultaneously, which rules out the operation of such a mechanism.

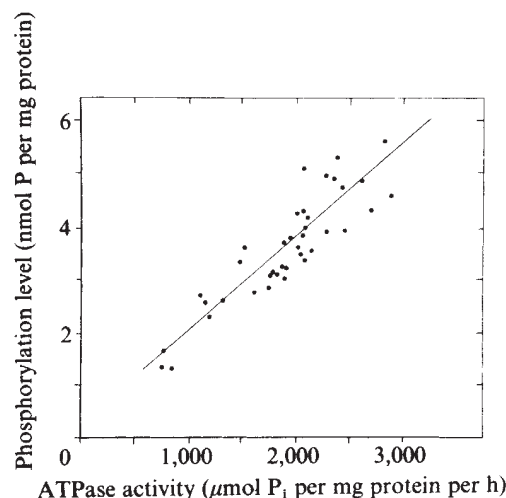
We have recently shown that the molar ratio of the  $\alpha$  and  $\beta$  subunits in the enzyme from the rabbit kidney is 1:1 (ref. 11). We have also determined the molecular weights of the subunits by sedimentation equilibrium analysis in the absence of detergent. After correction for carbohydrate and residual phospholipid contents, we obtained protein molecular weights of 120,600 for the  $\alpha$  subunit and 42,800 for the  $\beta$  subunit. This means that the smallest unit, containing the catalytic subunit,

would be an  $\alpha\beta$  dimer with a protein molecular weight of 163,400.

During these studies we noticed considerable discrepancy between results of the commonly used modified Lowry method<sup>12</sup> with bovine serum albumin as a standard and the absolute method by means of quantitative amino acid analysis. Recently, Craig and Kyte<sup>6</sup> have also indicated that the Lowry method may give falsely high results for this type of membrane protein. We have, therefore, compared the Lowry protein values for a number of  $(\text{Na}^+ + \text{K}^+)\text{ATPase}$  preparations with those from the absolute method of quantitative amino acid analysis in the presence of an internal standard (Table 1). The protein values obtained with the Lowry method are 1.345 (s.e.m. 0.018,  $n = 6$ ) times greater than those obtained with the absolute method.

This finding has far-reaching consequences for several parameters of  $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ , such as its turnover number and specific activity, and especially its phosphorylation capacity. Nearly all published values for these parameters are based on Lowry protein determinations according to the original procedure<sup>13</sup> or to modifications of this method; it would now seem that these values are generally too low.

We have also reinvestigated the phosphorylation capacity and found that the removal of endogenous ATP, which is added during purification of the enzyme to stabilize it, lowers the specific activity by 21% (s.e.m. 6,  $n = 11$ ). We have, therefore, determined the phosphorylation capacity both with and without prior removal of endogenous ATP. Results, plotted as phosphorylation capacity against specific activity for a given preparation, are shown in Fig. 1, with technical details in the legend. A linear relationship between these two parameters is found, with the values obtained with or without removal of ATP before phosphorylation falling along the same line. For the most active preparations a phosphorylation capacity of 5.6 nmol P per mg protein (based on amino acid analysis) is obtained, slightly



**Fig. 1** Relationship between phosphorylation capacity and specific activity of 41 purified  $(\text{Na}^+ + \text{K}^+)\text{ATPase}$  preparations from rabbit kidney. The enzyme was prepared as previously described<sup>12</sup>. Activity was determined according to Schoot *et al.*<sup>15</sup> and protein by quantitative amino acid analysis as described in Table 1 legend. Phosphorylation was carried out as described before<sup>16</sup> (34 preparations) or with the modification that the enzyme is phosphorylated without prior removal of endogenous ATP (seven preparations). In the latter case excess  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  (1–2.5 mM, diluted with cold ATP) was added. In both cases there is the same linear relationship between phosphorylation level and specific activity of  $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ . The equation for the regression line, calculated by the method of least squares, is  $y = 1.737 \cdot 10^{-3}x + 0.219$  ( $r = 0.90$ ).



**Table 1** Protein content of rabbit kidney ( $\text{Na}^+ + \text{K}^+$ )ATPase determined by the Lowry method and by amino acid analysis

| Preparation no. | Lowry method<br>(mg protein $\text{ml}^{-1}$ ) | Amino acid analysis<br>(a.a.a.) | Ratio<br>Lowry/a.a.a.      |
|-----------------|------------------------------------------------|---------------------------------|----------------------------|
| 1               | 4.80                                           | 3.67                            | 1.308                      |
| 2               | 1.52                                           | 1.14                            | 1.333                      |
| 3               | 2.01                                           | 1.47                            | 1.367                      |
| 4               | 2.03                                           | 1.48                            | 1.372                      |
| 5               | 1.67                                           | 1.19                            | 1.403                      |
| 6               | 2.05                                           | 1.59                            | 1.289                      |
| Mean            |                                                |                                 | $1.345 \pm 0.018$ (s.e.m.) |

A 100- $\mu\text{l}$  aliquot of a ( $\text{Na}^+ + \text{K}^+$ )ATPase suspension ( $\sim 2$  mg protein  $\text{ml}^{-1}$ ) in 0.25 M sucrose, 10 mM Tris-HCl, 1 mM EDTA (pH 7.4) was added to a glass tube containing 150 nmol norleucine, followed by 300  $\mu\text{l}$  7.6 M HCl. The tube was vacuum sealed and hydrolysed in the dark for 24 h at 105 °C. After 10 min centrifugation at 20,000g, amino acid analysis of the supernatant was carried out in a modified JEOL type JLC-6AH amino acid analyser. The results of the amino acid analysis were corrected for incomplete hydrolysis of valine and isoleucine and degradation of serine and tryptophan<sup>11</sup>. The protein content was calculated by summation of the amino acid weights. Parallel samples of the ( $\text{Na}^+ + \text{K}^+$ )ATPase suspension were analysed for protein according to Lowry *et al.*<sup>13</sup> after trichloroacetic acid precipitation as described by Jørgensen<sup>12</sup>.

higher than the recently reported<sup>14</sup> value of 5.3 nmol P per mg protein.

The maximal phosphorylation capacity of 5.6 nmol P per mg protein at an  $\alpha\beta$  dimer protein molecular weight of 163,400 corresponds to 0.92 phosphorylation site per  $\alpha\beta$  dimer. As even a highly purified enzyme preparation is not 100% pure on a protein basis (as seen by the appearance of minor protein bands beyond the two prominent bands belonging to the  $\alpha$  and  $\beta$  subunits on SDS-gel electrophoresis, data not shown), this value is a slight underestimation of the true phosphorylation capacity.

Thus we conclude that every  $\alpha\beta$  unit has one active phosphorylation site. As the molar binding values for ATP, ouabain and vanadate equal the phosphorylation level<sup>1,7</sup>, our results also indicate that each  $\alpha\beta$  unit can bind one molecule each of ATP, ouabain and vanadate. As the most likely subunit composition of the enzyme is  $\alpha_2\beta_2$  (ref. 11), we may safely conclude that the  $\alpha_2\beta_2$  tetramer has two phosphorylation sites, which can be phosphorylated simultaneously. This rules out the operation of a 'half-of-the-sites mechanism' in ( $\text{Na}^+ + \text{K}^+$ )ATPase from rabbit kidney.

We thank Annie M. M. Fleuren-Jakobs for technical assistance.

## Diversity of immunoglobulin expression in leukaemic cells resembling B-lymphocyte precursors

L. B. Vogler, J. L. Preud'homme & M. Seligmann

Laboratory of Immunochemistry and Immunopathology (INSERM U108), Institut de Recherches sur les Maladies du Sang de l'Université Paris VII; Laboratoire d'Oncologie et d'Immuno-Hématologie du CNRS, Hôpital Saint Louis, 75475 Paris Cedex 10, France

W. E. Gathings, W. M. Crist & M. D. Cooper

Cellular Immunobiology Unit, Department of Pediatrics and Microbiology and the Comprehensive Cancer Center, University of Alabama in Birmingham, Birmingham, Alabama 35294, USA

F. J. Bollum

Department of Biochemistry, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20014, USA

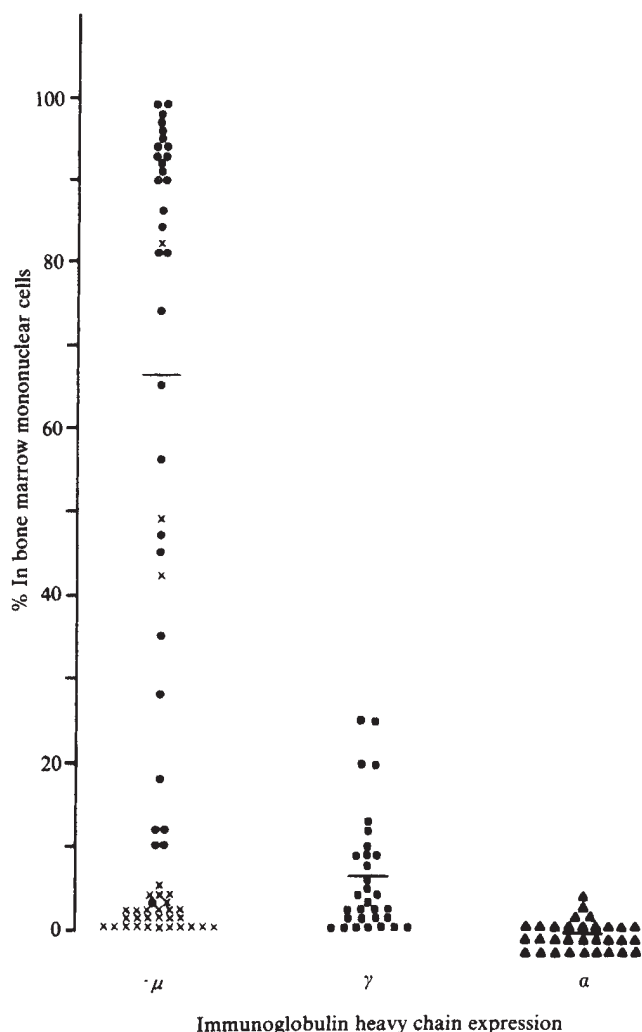
Approximately 20% of patients with acute lymphocytic leukaemia (ALL) have leukaemic blasts with features of pre-B cells<sup>1</sup> which are the recently characterized precursors of B lymphocytes in normal development (for a review, see ref. 2). Pre-B cells isolated from normal bone marrow or fetal liver, and malignant cells from patients with pre-B cell leukaemia, are rapidly dividing lymphoid cells that contain cytoplasmic immunoglobulin  $\mu$  heavy chains, but have no detectable surface immunoglobulin. The resemblance of immunoglobulin-containing ALL cells to normal precursors of B lymphocytes and their availability in relatively pure preparations allowed us to explore them as models of early stages in the differentiation of the B-lymphocyte line. We report here observations on the occurrence of intermediate pre-B/B-cell phenotypes, immunoglobulin isotype switching and the asynchrony of immunoglobulin heavy and light chain expression in 30 cases of ALL and 3 cases of chronic myelogenous leukaemia in lymphoblastic crisis (CML-BC).

Heparinized bone marrow from patients with untreated ALL or with CML-BC was examined after lysis of erythrocytes with a hypotonic solution of ammonium chloride. Surface immunoglobulins were detected by immunofluorescent staining of fresh, viable cells with fluorochrome-conjugated, affinity-purified goat anti-human  $\mu$ ,  $\gamma$ ,  $\alpha$ ,  $\kappa$  and  $\lambda$  antibodies, as described elsewhere<sup>3</sup>. Intracytoplasmic immunoglobulin was detected by immunofluorescent staining of cells that had been centrifuged or smeared on to glass slides, fixed in cold 95% ethanol, 5% acetic acid and washed in buffered saline as described elsewhere<sup>1</sup>. Differently coloured fluorochromes (rhodamine and fluorescein isothiocyanates) were used to stain cells before and after fixation, which allowed us to distinguish between surface and cytoplasmic determinants in single cells. In some patients observations were reproduced using a completely different set of reagents: fluorochrome-conjugated F(ab')<sub>2</sub> fragments of rabbit anti-human  $\mu$ ,  $\gamma$ ,  $\alpha$ ,  $\kappa$  and  $\lambda$  antibodies as described elsewhere<sup>4</sup>. Terminal deoxynucleotidyl transferase was detected by indirect immunofluorescence using a rabbit anti-terminal deoxynucleotidyl transferase serum<sup>5</sup> and fluorescein-conjugated goat anti-rabbit Fab. Results of immunofluorescence studies represent concurrent observations of at least two of the authors.

In 1 year we found 33 patients with intracytoplasmic immunoglobulin-positive lymphoblastic leukaemia. In 31 of these at least 10% of marrow blasts contained cytoplasmic and no surface immunoglobulin, as originally observed in pre-B

Received 4 November 1980; accepted 9 February 1981.

- Jørgensen, P. L. *Physiol. Rev.* **60**, 864-917 (1980).
- Peterson, G. L., Ewing, R. D., Hootman, S. R. & Conte, F. P. *J. biol. Chem.* **253**, 4762-4770 (1978).
- Perrone, J. R., Hackney, J. F., Dixon, J. F. & Hokin, L. E. *J. biol. Chem.* **250**, 4178-4184 (1975).
- Esmann, M., Skou, J. C. & Christiansen, C. *Biochim. biophys. Acta* **567**, 410-420 (1979).
- Hastings, D. F. & Reynolds, J. A. *Biochemistry* **18**, 817-821 (1979).
- Craig, W. S. & Kyte, J. J. *J. biol. Chem.* **255**, 6262-6269 (1980).
- Hansen, O., Jensen, J., Nørby, J. G. & Ottolenghi, P. *Nature* **280**, 410-412 (1979).
- Askari, A. & Huang, W. *Biochem. biophys. Res. Commun.* **93**, 448-453 (1980).
- Stein, W. D., Lieb, W. R., Karlisch, S. J. D. & Eilan, Y. *Proc. natn. Acad. Sci. U.S.A.* **70**, 275-278 (1973).
- Repke, K. R. H. & Schön, R. *Acta biol. med. germ.* **31**, K19-K30 (1973).
- Peters, W. H. M., De Pont, J. J. H. H. M., Koppers, A. & Bonting, S. L. *Biochim. biophys. Acta* **641**, 55-71 (1981).
- Jørgensen, P. L. *Biochim. biophys. Acta* **356**, 36-52 (1974).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
- Askari, A., Huang, W. & Antieau, J. M. *Biochemistry* **19**, 1132-1140 (1980).
- Schoot, B. M., Schoots, A. F. M., De Pont, J. J. H. H. M., Schuurmans Stekhoven, F. M. A. H. & Bonting, S. L. *Biochim. biophys. Acta* **483**, 181-192 (1977).
- Schuurmans Stekhoven, F. M. A. H., Swarts, H. G. P., De Pont, J. J. H. H. M. & Bonting, S. L. *Biochim. biophys. Acta* **597**, 100-111 (1980).



**Fig. 1** Surface and cytoplasmic expression of immunoglobulin isotypes in leukaemic cells. ×, Surface  $\mu$ ; ●, ■, ▲, cytoplasmic  $\mu$ ,  $\gamma$  and  $\gamma$  chains, respectively. Horizontal bars represent the mean percentage of marrow mononuclear cells expressing intracytoplasmic immunoglobulin of each class.

leukaemia. In contrast, normal pre-B cells comprise only  $1.0 \pm 0.6\%$  of human marrow mononuclear cells<sup>6</sup>. When taken collectively, blasts of pre-B phenotype accounted for 66.7% of the leukaemic cells in these patients (Fig. 1). In individuals, however, leukaemic cells of identical morphology were sometimes found to be composed of three distinct populations based on immunoglobulin expression: cells showing no immunoglobulin, cells with intracytoplasmic immunoglobulin only, and cells with both intracytoplasmic and scant amounts of surface  $\mu$ . In three patients cells of the latter type accounted for a large proportion (42, 49 and 82%) of the lymphoblasts (Fig. 1). This phenotype, also observed by Brouet *et al.*<sup>7</sup> and Greaves *et al.*<sup>8</sup>, is thought to represent cells at a stage of differentiation intermediate between that of the pre-B and B-cell stages. Several features distinguish these 'intermediate' cells from mature B lymphocytes: they have persistent expression of cytoplasmic  $\mu$  chains and the nuclear enzyme, terminal deoxynucleotidyl transferase; and in two of three cases the sparse amounts of  $\mu$  chain on the cell surface were not associated with detectable light chains.

As in previous studies<sup>1,7</sup>, many patients with pre-B leukaemia had a very small proportion of a leukaemic cells that expressed

cytoplasmic  $\gamma$  or  $\alpha$  chains in addition to cytoplasmic  $\mu$ . However, in five of our present patients, cells containing  $\gamma$  chains accounted for a relatively large proportion (>10%) of the leukaemic population, sometimes outnumbering  $\mu$ -containing leukaemic cells (Fig. 2, Table 1).

In addition to the 31 patients already described, one with ALL and one with CML-BC had a small proportion of lymphoblasts (8% and 6%, respectively) that contained only  $\gamma$  heavy chains. Unlike the previous cases,  $\mu$  chains (and other non- $\gamma$  chains) were undetectable in both the  $\gamma$ -positive and  $\gamma$ -negative lymphoblasts from these individuals. As  $\gamma$ -containing cells comprised <10% of mononuclear cells in the marrow of these patients and  $\mu$  chains were not detected, they did not conform to our original definition of pre-B cell leukaemia. However, there is evidence for the immaturity of these leukaemic cells and for their distinction from IgG-secreting lymphoplasmablasts: (1) the distribution and intensity of immunofluorescent staining with a pattern identical to that of  $\mu$ -staining in pre-B cells; (2) the co-expression of cytoplasmic  $\gamma$  and nuclear terminal deoxynucleotidyl transferase in double immunofluorescence studies; and (3) the absence of detectable light chains in freshly collected lymphoblasts. Note also that  $\kappa$  chains were found in 10% of the  $\gamma$ -positive lymphoblasts from the patient with CML-BC when leukaemic cells were cultured for 24 h in RPMI 1640 medium supplemented with fetal calf serum.

Cells from 15 patients were examined for surface and cytoplasmic  $\kappa$  and  $\lambda$  light chains (Table 1). In only four of these were cytoplasmic light chains found in most of leukaemic cells containing  $\mu$  heavy chains. Light chains usually were present in only a small proportion of cells containing heavy chains, or were undetectable. In five patients (see Table 1, patients 1, 2, 4, 10 and 11) when leukaemic cells were examined simultaneously for light and heavy chains by double immunofluorescent staining, the expression of cytoplasmic  $\kappa$  chains was associated with that of  $\gamma$  heavy chains. No surface light chains were detected in the absence of surface heavy chains.

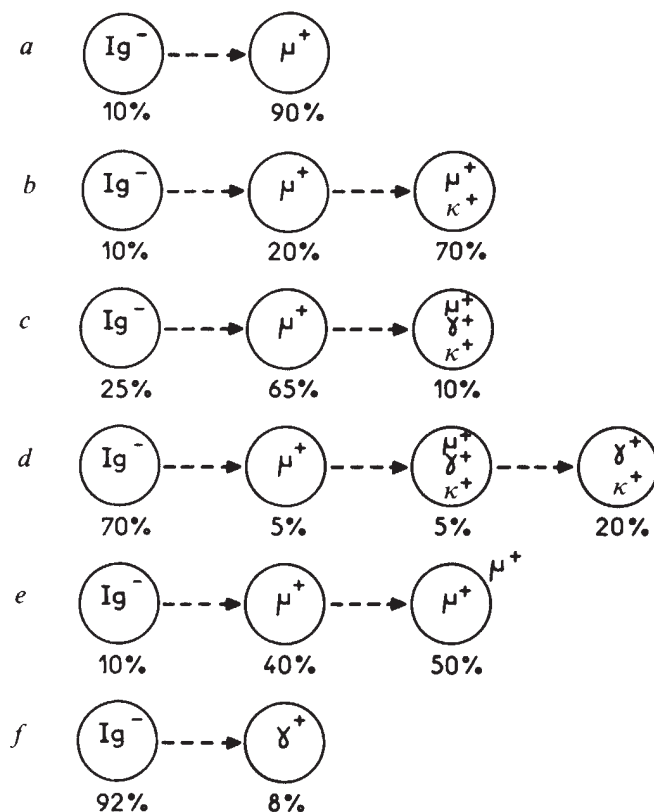
Analysis of leukaemic cells from our patients with ALL and CML-BC reveals that immunoglobulin expression in the pre-B group of leukaemias is less homogeneous than originally thought (Fig. 2). Cells carrying sparse surface  $\mu$ , probably representing an intermediate pre-B/B-cell stage of development, were predominant in three patients. The infrequent expression of light chains relative to  $\mu$  chains in pre-B leukaemic cells from most of these patients agrees with similar results reported by Brouet *et al.*<sup>7</sup> and Greaves *et al.*<sup>9</sup>. An asynchrony in the expression of cytoplasmic  $\mu$  heavy chains and light chains has also recently been observed in murine cells of pre-B phenotype from normal fetal liver<sup>10</sup>, Abelson virus-infected bone

**Table 1** Cytoplasmic immunoglobulin expression by lymphoblasts from patients with ALL or CML-BC

| Patient | Diagnosis | Ig heavy chain |          |          | Ig light chain |           |
|---------|-----------|----------------|----------|----------|----------------|-----------|
|         |           | $\mu$          | $\gamma$ | $\alpha$ | $\kappa$       | $\lambda$ |
| 1       | ALL       | 10             | 25       | <1       | 25             | <1        |
| 2       | ALL       | 81             | 10       | <1       | 15             | <1        |
| 3       | ALL       | 92             | <1       | <1       | 5              | <1        |
| 4       | ALL       | 36             | 13       | <1       | 14             | <1        |
| 5       | ALL       | 99             | 5        | <1       | <1             | <1        |
| 6       | ALL       | 95             | 3        | <1       | <1             | <1        |
| 7       | ALL       | 93             | 4        | 2        | 3              | 3         |
| 8       | ALL       | 90             | <1       | <1       | 70             | <1        |
| 9       | ALL       | 86             | 4        | <1       | 87             | <1        |
| 10      | ALL       | 65             | 25       | <1       | 25             | <1        |
| 11      | ALL       | 10             | 20       | <1       | 20             | <1        |
| 12      | CML-BC    | 81             | <1       | <1       | <1             | <1        |
| 13      | CML-BC    | 3              | 9        | <1       | 2              | <1        |
| 14      | ALL       | <1             | 8        | 1        | 2              | <1        |
| 15      | CML-BC    | <1             | 6        | <1       | 1              | <1        |

Results are expressed as percentages of positive cells among bone marrow mononuclear cells.





**Fig. 2** Phenotypes observed in immunoglobulin-containing lymphoblastic leukaemias. Six patients were selected to demonstrate the heterogeneity of immunoglobulin expression in immunoglobulin-positive lymphoblastic leukaemias. A representative cell phenotype comprises at least 5% of the leukaemic cell population. In each case a varied number of leukaemic blasts expressed no detectable immunoglobulin and could be described as 'null' phenotype. The surface μ<sup>+</sup> leukaemic cells in patient *e* may represent a developmental stage intermediate between that of pre-B and B cell. Both these cells and the γ<sup>+</sup> cells in patient *f* had no detectable light chains, but contained nuclear terminal deoxynucleotidyl transferase. Dashed arrows show suggested directions of differentiation.

marrow-derived cell lines<sup>11</sup>, and pre-B/myeloma cell hybridomas<sup>12</sup>. In humans the study of bone marrow and fetal liver also indicates that only a small minority (<10%) of normal pre-B cells co-express κ or λ light chains with μ heavy chains; most normal human pre-B cells express only μ (H. Kubagawa, unpublished results). In contrast to these results and observations in most patients with pre-B leukaemia, κ light chains were detected in >5% of leukaemic cells from seven of our patients, usually in μ-containing cells, and their incidence coincided with the expression of γ chains in five patients. However, the absence of γ chains in cytoplasmic μ<sup>+</sup>, κ<sup>+</sup> cells from two patients, and the absence of light chains in γ<sup>+</sup> cells from two others, precludes any simple explanations for γ and light-chain concurrence during differentiation.

The expression of a second heavy chain isotype in some of our patients' leukaemic cells has not been previously observed in normal human pre-B cells. There are two possible explanations for this: because γ chains were observed in a small population of leukaemic cells from a minority of patients with pre-B leukaemia, class switching may be a rare event at the pre-B stage of differentiation. Therefore, it would be difficult to detect a small population of γ-containing cells among the ~1% of μ-positive pre-B cells present in normal marrow using conventional tech-

niques, although we have made considerable efforts in this direction. Alternatively, the γ chains found in these leukaemic cells may reflect abnormal regulation of immunoglobulin gene expression related to malignant transformation, rather than the sequential events of normal B-lymphocyte development.

Whichever the case, our results indicate that the mechanisms responsible for the isotype switch from μ to γ chains and for light chain gene expression are operable at a stage in B-cell development before the expression of surface immunoglobulin and antigen receptors. We suggest that cells containing cytoplasmic γ chains and with the phenotypes described here, reflect early stages in the development of normal B cells, perhaps representing pathways alternative to the major sequence of maturation.

We thank Drs G. Flandrin, D. Buriot, G. Duhamel of Paris, and members of the Southwest Oncology Group, USA, for contributing samples for this study, Dr H. Kubagawa and Mrs S. Labaume for preparing antibodies and Mrs K. Nelson for technical assistance. This work was supported by US PHS grants CA 25408, CA 16673, CA 13148 and DGRST grants 80.7.0220, 78.7.3101. L.B.V. is a Leukemia Society of America Special Fellow and an INSERM Foreign Research Associate.

Received 27 November 1980; accepted 19 January 1981.

1. Vogler, L. B. *et al.* *New Engl. J. Med.* **298**, 872-878 (1978).
2. Cooper, M. D. & Lawton, A. R. in *Cells of Immunoglobulin Synthesis* (eds Pernis, B. & Vogel, H. J.) 411-425 (Academic, New York, 1979).
3. Gathings, W. E., Lawton, A. R. & Cooper, M. D. *Eur. J. Immun.* **7**, 804-810 (1977).
4. Preud'homme, J. L. & Labaume, S. in *In Vitro Methods in Cell-Mediated and Tumor Immunity* (eds Bloom, B. R. & David, J. J.) 155-169 (Academic, New York, 1976).
5. Bollum, F. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4119-4122 (1975).
6. Pearl, E. R. *et al.* *J. Immun.* **120**, 1169-1175 (1975).
7. Brouet, J. C. *et al.* *Blood* **54**, 269-273 (1979).
8. Greaves, M. F. *et al.* *Leukemia Res.* **3**, 353-362 (1979).
9. Greaves, M. F. *et al.* *Leukemia Res.* **3**, 181-191 (1979).
10. Levitt, D. & Cooper, M. D. *Cell* **19**, 617-625 (1980).
11. Siden, E. J., Baltimore, D., Clark, D. & Rosenberg, N. E. *Cell* **16**, 389-396 (1979).
12. Burrows, P., Lejeune, M. & Kearney, J. F. *Nature* **280**, 838-840 (1979).

## Natural occurrence but lack of melanotrophic activity of γ-MSH in fish

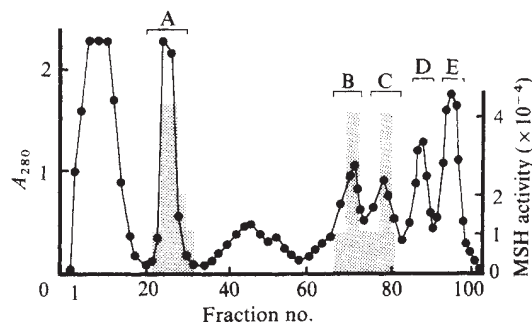
C. McLean & P. J. Lowry

Pituitary Hormone Laboratory, Department of Chemical Pathology, St Bartholomew's Hospital, London EC1A 7BE, UK

When adrenocorticotropin (ACTH) and β-lipotropin (LPH) are enzymatically cleaved in the pituitary from their common precursor molecule, a glycopeptide with an apparent molecular weight of 16,000 remains (reviewed in ref. 1). Sequencing of a cDNA copy of the bovine pituitary mRNA for the precursor molecule demonstrated that, within the region encoding the 16,000-molecular weight (MW) glycopeptide, there was a short sequence homologous with the α-melanocyte-stimulating hormone (α-MSH) contained in ACTH and the β-MSH within β-LPH<sup>2</sup>. It was therefore proposed that the region of homology coded for a γ-MSH which might have biological activity, although synthetic peptides containing this sequence were shown to have poor melanotrophic activity<sup>3</sup>. In the pars intermedia further rapid enzyme modification converts ACTH to α-MSH and corticotropin-like intermediate lobe peptide (CLIP), and β-LPH to β-MSH and β-endorphin. Although smaller peptides containing the γ-MSH structure have been detected in extracts of the bovine pars intermedia<sup>4,5</sup>, the major

post-translational products of the 16,000-MW fragment are present mainly as large glycosylated peptides<sup>1,4</sup>. We report here that a major peptide product of the neurointermediate lobe of the dogfish pituitary which we had previously shown to lack melanotrophic activity has a  $\gamma$ -MSH-like structure. Therefore if  $\gamma$ -MSH is a peptide with a function it is not one shared with the melanotrophic hormones.

In our previous studies, peptides analogous to mammalian  $\alpha$ - and  $\beta$ -MSH were characterized from extracts of the neurointermediate lobe of a more phylogenetically ancient species, the dogfish (*Squalus acanthias*)<sup>6</sup>, the final purification of which is shown in Fig. 1. Peak A was identified as dogfish  $\beta$ -MSH



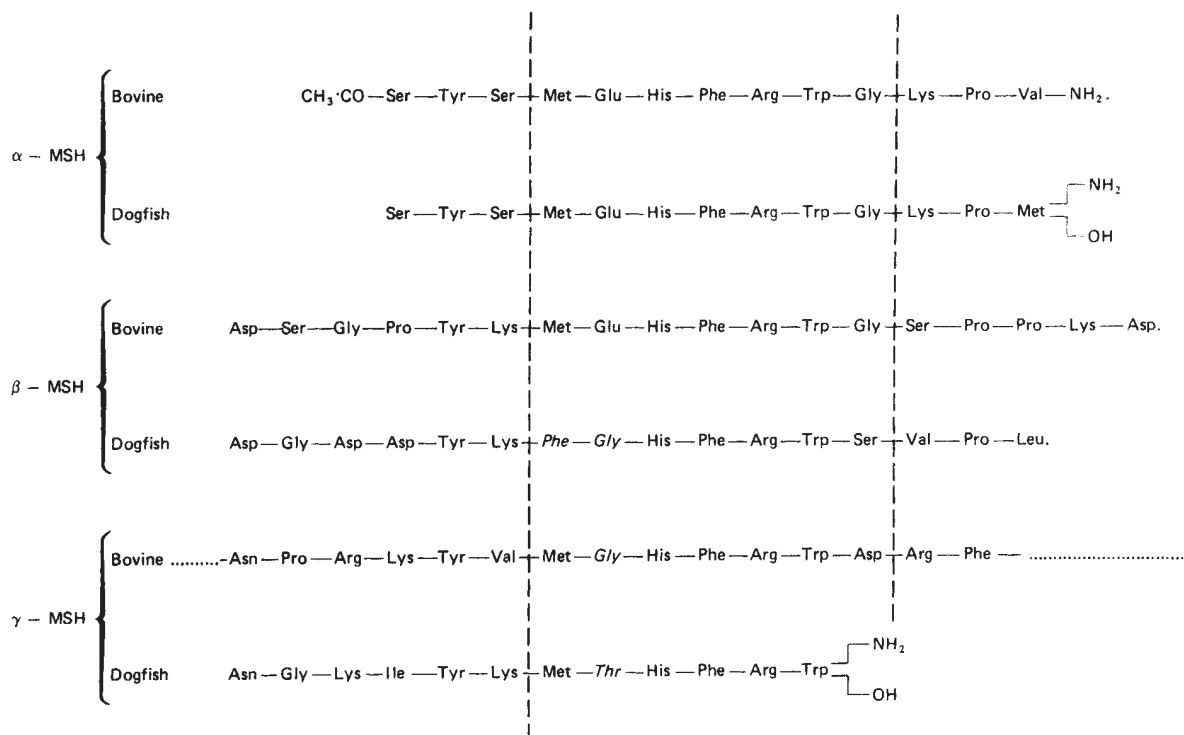
**Fig. 1**  $A_{280}$ ; MSH activity is shown by the hatched areas. The column (0.6 cm i.d.  $\times$  25 cm length) was developed with a concave gradient running from 20 to 500 mM trimethylamine acetate, pH 5 (total volume 150 ml), and 1-ml fractions were collected every 20 min. A, B and C denote peaks of MSH activity. (From ref. 6.)

whilst B and C were two forms of  $\alpha$ -MSH (C was found to be the C-terminally amidated form of B). At concentrations where peaks A and B showed potent MSH activity, D and E were inactive.

Peaks D and E stored from the previous study were repurified by preparative high-pressure liquid chromatography (HPLC,  $1 \times 10$  cm Spherisorb ODS  $10 \mu\text{m}$ ) developed with a methanol gradient (40–80% in  $\text{H}_2\text{O}$ , containing 1% trifluoroacetic acid). Amino acid analysis of samples of each purified peptide after both acid and enzyme hydrolysis gave identical compositions: Asn<sub>1</sub>, Thr<sub>1</sub>, Gly<sub>1</sub>, Met<sub>1</sub>, Ile<sub>1</sub>, Tyr<sub>1</sub>, Phe<sub>1</sub>, His<sub>1</sub>, Lys<sub>2</sub>, Trp<sub>1</sub> and Arg<sub>1</sub>. E was deduced to be a C-terminally amidated form of D in view of its higher ammonia values and its more basic behaviour during purification on CM-cellulose (Fig. 1). D (50  $\mu\text{g}$ ) was submitted to automated sequence analysis and gave the unambiguous structure shown in Fig. 2, which also shows the equivalent sequence predicted from the bovine cDNA study and comparisons of the other two species of melanotrophins.

Although there is little sequence homology between D (and its amidated form E) and the  $\gamma$ -MSH bovine sequence, their lack of melanotrophic activity compared with the two other dogfish peptides,  $\alpha$ - and  $\beta$ -MSH, and the poor melanotrophic activity of synthetic bovine  $\gamma$ -MSH peptides<sup>3</sup> suggest that their assignment as  $\gamma$ -MSH peptides is correct. It is interesting to note here that the unconservative substitutions in the core sequence of dogfish  $\beta$ -MSH compared with bovine  $\beta$ -MSH (Phe for Met and Gly for Glu) have maintained its potent MSH activity whilst both  $\gamma$ -MSHs which have a more recognizable core sequence (Met-X-His-Phe-Arg-Trp) are comparatively inactive as melanotrophic peptides.

The poor periodic acid-Schiff staining of the dogfish (and other fishes, for example, teleosts) intermediate tissue<sup>8</sup> suggests that the fish  $\gamma$ -MSH/ACTH/LPH precursor is not glycosylated,



**Fig. 2** Comparison of the known and predicted sequences of bovine and dogfish MSH peptides. Sequence analysis of dogfish  $\gamma$ -MSH was carried out in the JEOL-47K sequence analyser using 3 mg Polybrene as carrier. The normal protein programme was used with 0.5 M Quadrol as the reaction buffer. The nitrogen-dried phenyl thiazolinone derivatives were converted to their phenylthiohydantoin derivatives by treatment with 100  $\mu\text{l}$  0.5 M HCl at 70 °C for 10 min. After addition of 20  $\mu\text{l}$  1 M trisodium citrate the derivatives were directly analysed by HPLC (Altex-ODS, 10–40% acetonitrile/tetrahydrofuran in 10 mM triethylamine phosphate, pH 3, 20 min). Confirmation of samples was carried out by amino analysis after back hydrolysis with HI (150 °C, 24 h).



as in the mammals, near the  $\gamma$ -MSH region. It may be this lack of glycosylation that allows the precursor to be processed to the two  $\gamma$ -MSH-like peptides in fish but not in mammals. Their occurrence as major peptide products of the pars intermedia and the C-terminal amidation suggest a biological role for these peptides in the fish but it is unfortunate that the term  $\gamma$ -MSH was coined for such peptides when they have little if any, melanotrophic activity in an animal, like the dogfish, which responds to changes in background by altering skin pigmentation.

We thank the *Biochemical Journal* for permission to reproduce Fig. 1 and Dr J. P. Sumpter for comments.

Received 27 October 1980; accepted 15 January 1981.

1. Eipper, B. A. & Mains, R. E. *Endocr. Rev.* **1**, 1-27 (1980).
2. Nakanishi et al. *Nature* **278**, 423-427 (1979).
3. Ling, N. et al. *Life Sci.* **25**, 1773-1780 (1979).
4. Shibasaki, T. et al. *Nature* **285**, 416-417 (1980).
5. Shibasaki, T. et al. *Biochem. biophys. Res. Commun.* **96**, 1393-1399 (1980).
6. Bennett, H. P. J. et al. *Biochem. J.* **141**, 439-444 (1974).
7. Bennett, H. P. J. et al. *Biochem. J.* **129**, 695-701 (1972).
8. Holmes, R. L. & Ball, J. N. (eds) *Biological Structure and Function. The Pituitary Gland: a Comparative Account* Vol. 4 (Cambridge University Press, 1974).

## Slow-reacting substances (leukotrienes) contract human airway and pulmonary vascular smooth muscle *in vitro*

C. J. Hanna\*, M. K. Bach†, P. D. Pare\* & R. R. Schellenberg\*

\* University of British Columbia, Pulmonary Research Laboratory, 1081 Burrard Street, Vancouver, British Columbia, Canada V6Z 1Y6

† Department of Hypersensitivity Disease Research, The Upjohn Company, Kalamazoo, Michigan 49001, USA

During a type I allergic reaction histamine, slow-reacting substance of anaphylaxis (SRS-A) and other mediator substances are elaborated from specific tissue sites<sup>1</sup>. In allergic asthma these sites are in the lung and the mediator substances cause airway obstruction by contracting smooth muscle and altering mucociliary function<sup>2-4</sup>. Unlike histamine, slow-reacting substances (SRSs) have been assessed very little for their roles in obstructive airways disease. This has been partly due to the fact that their chemical nature was unknown until recently and thus pure samples were not available for pharmacological studies. However, SRSs isolated from both immunological and non-immunological reactions have been identified as a combination of two related lipid substances—leukotriene C<sub>4</sub> (LTC) and leukotriene D<sub>4</sub> (LTD)<sup>5,6</sup>; thus it is now possible to use pure SRSs (leukotrienes) in pharmacological studies of airway smooth muscle. LTC and LTD have been shown to contract guinea pig tracheal and lung parenchymal strips<sup>5</sup> but there is no evidence that these substances produce similar effects on human lung tissue. To clarify this, *in vitro* pharmacological studies were done to determine the actions of LTC and LTD on smooth muscle strips of human bronchus, pulmonary vein and artery, and lung parenchymal tissue containing smooth muscle components and pleura. As indicated in a preliminary report<sup>7</sup>, all four types of tissues contracted in a dose-dependent fashion to the leukotrienes, although these substances only function as partial agonists.

Lung tissues were obtained from 14 patients undergoing lobectomy or pneumonectomy for isolated peripheral pulmonary lesions. Samples of bronchus, pulmonary vein, pulmonary artery and parenchyma were dissected free and placed in oxygenated Krebs' solution. The bronchial, pulmonary venous and pulmonary arterial samples were then cut into spiral strips 1×4×40 mm; parenchymal strips, with and without pleura, were cut in strips about 5×5×40 mm. These were placed in an organ bath system containing Krebs' solution and drug responses were examined. SRS peaks I and II were obtained from rat peritoneal monocytes by methods described elsewhere<sup>6,8</sup>. Peak I was identified as leukotriene D<sub>4</sub> (ref. 9) and peak II as leukotriene C<sub>4</sub> (formerly designated LTC-1)<sup>10</sup>.

A cumulative drug application procedure<sup>11</sup> was used to study the contractile effects of all the agonists, as well as the antagonistic action of FPL 55712 against the leukotrienes, on the four types of lung tissue. All drug concentrations are expressed in molar units. LTD and LTC molarities were determined by converting Upjohn SRS units of activity into moles of leukotriene (LTD: 1 U = 20 pg, molecular weight (MW) = 495; LTC: 1 U = 24±2 pg, MW = 625) and dividing by the dilution volume in litres.

The log dose-response curves for carbachol, histamine (both Sigma), LTC and LTD on human bronchial smooth muscle are shown in Fig. 1. Carbachol and histamine elicited fast onset, sustained, dose-dependent contractions, although the bronchial tissue was consistently more sensitive to carbachol over the full range of the dose-response curve. Carbachol also produced the greatest magnitude of contraction compared to the other agonists. Both carbachol and histamine were active over a wide dose range ( $5 \times 10^{-9}$ – $5.5 \times 10^{-4}$  M) with  $pD_2$  values (calculated as  $-\log ED_{50}$ ) of 6.301 and 5.658, respectively. Submaximal effects of both agonists were easily reversed when they were washed from the organ baths.

However, responses to the leukotrienes were markedly different: both LTC and LTD caused human bronchial smooth muscle to contract but the responses had a delayed onset and slow development time. The contractions were well sustained (that is, they did not fade) and were dose-dependent. As can be seen in Fig. 1, the bronchial tissue was most sensitive to LTC although this leukotriene produced maximum responses that were only 60–70% that of the maximum responses to carbachol.

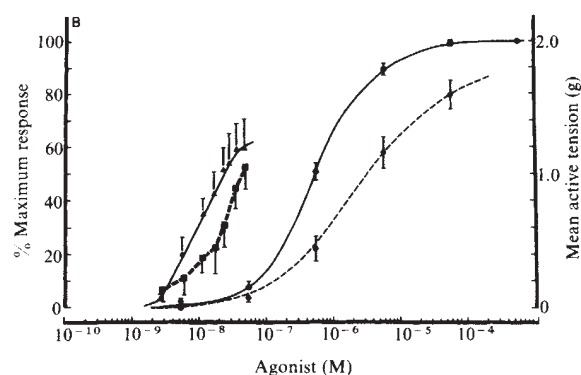
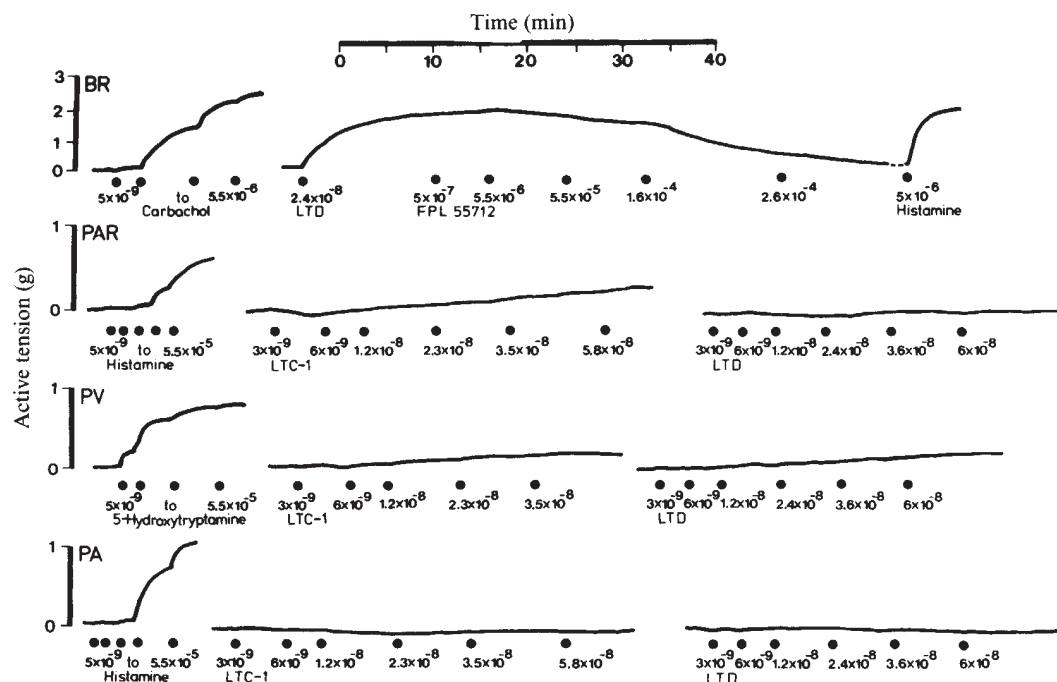


Fig. 1 Log dose-response curves for carbachol (●), histamine (◆), LTC (▲) and LTD (■) on human bronchial smooth muscle. Tissues obtained from patients undergoing thoracic surgery for carcinoma were cut into spiral strips and maintained in 20 ml jacketed organ baths containing oxygenated Krebs' solution, pH 7.4, at 37 °C. Contractions were measured semi-isometrically with Grass FT .03 C force displacement transducers and recorded on a Beckman R511A polygraph. Initial tensions ranged from 1 to 2 g; raw data were collected as active tension (g). Data points are expressed as a percentage of the maximum carbachol response (left ordinate) and a mean active tension scale (right ordinate) is shown for comparison. Each point represents the mean ± s.e.m. of three to seven experiments.



**Fig. 2** Original polygraph tracings of the responses of human bronchus (BR), parenchyma (PAR), pulmonary vein (PV) and pulmonary artery (PA) to selected agonists and the antagonist, FPL 55712. Contractile responses are shown in terms of active tension (g) generated per unit of time (min). Drugs were added to the bath in a cumulative manner. LTD, leukotriene  $D_4$ ; LTC, leukotriene  $C_4$ .

LTD similarly acted as a partial agonist. Both the leukotrienes were active over a narrow dose range ( $3 \times 10^{-9}$ – $5 \times 10^{-8}$  M) with  $pD_2$  values of 8.022 for LTC and 7.620 for LTD. Human bronchial tissue was more variable in its responsiveness to LTC and LTD than to carbachol or histamine. It was also found that leukotriene-induced contractions were reversed when these substances were washed from the baths. Furthermore, exposure of tissues to leukotrienes (followed by washing) before addition of another agonist did not alter the tissue responses to the second agonist.

The SRS-A antagonist, FPL 55712 (Fisons), inhibited LTC- and LTD-induced contractions of human bronchial smooth muscle. The upper trace in Fig. 2 shows the response of a human bronchial strip to LTD ( $600 \text{ U} = 2.4 \times 10^{-8} \text{ M}$ ) followed by the addition of FPL 55712. The rate of inhibition was slow and the concentrations of FPL 55712 required to abolish completely the LTD contraction were high ( $2.6 \times 10^{-4} \text{ M}$ ), although the antagonism was clearly dose-dependent. No specific anti-histamine effect was observed with high doses of FPL 55712; the antagonist alone had no effect on the tissues.

The contractile effects of the leukotrienes and other selected agonists on human parenchymal tissue, pulmonary vein and pulmonary artery were also examined to a limited extent. Examples of traces are shown in Fig. 2. In general, the vasoactive amines, histamine or 5-hydroxytryptamine, were more active than LTC or LTD on each of these three types of lung tissue. The parenchymal tissue was slightly more sensitive to LTC (threshold 150 U) than LTD (300 U) and gave a slightly greater maximal contraction (LTC 60%, LTD 35% of the histamine maximum). Pulmonary vein and artery were not as responsive to the leukotrienes as the airway tissues and the responses were characteristically slow in onset and development compared with the amines.

These results suggest that LTC and LTD may act as airway spasmogens when released by an allergic reaction in the human lung. Our data show that human bronchial smooth muscle was very sensitive to the leukotrienes although both substances acted as partial agonists compared with carbachol. Previous studies<sup>12–14</sup> have shown that human bronchial smooth muscle contracted in response to crude preparations of SRSs; however, our results show that the distinctive 'slow reaction' is caused by the leukotrienes. Dahlen *et al.*<sup>15</sup> have described similar findings with LTC and LTD.

The inability of the leukotrienes to elicit a maximum contractile response, the delayed onset of contraction and the slow rate of FPL 55712 antagonism suggest a mechanism of action on human lung smooth muscle which differs from that seen for other agonists. Such differences indicate that the mechanism of action of the leukotrienes is not closely linked to the contractile apparatus.

In these studies, neither LTC nor LTD had preferential activity on central (bronchial) versus peripheral (parenchymal) airway smooth muscle. As the parenchymal tissues obtained were from patients with usually significant emphysema, the reactivity varied considerably from one tissue to another and we may thus have limited our ability to see a preferential peripheral response. However, this seems unlikely because unreactive tissues failed to respond to all agonists, whereas reactive parenchymal strips not only contracted to agonists but in the same approximate dose range as for the bronchial strips.

These results suggest the possibility that leukotrienes C and D may play a major part in human bronchial asthma.

Histamine and carbachol were from Sigma and FPL 55712 was from Fisons. We acknowledge the support of the Canadian and British Columbia Lung Associations and B.C. Health Care Research Foundation. We thank Drs J. C. Hogg and J. M. Nelems for their contributions and Mrs L. Hare for typing the manuscript. C.J.H. is a Canadian Lung Association Fellow.

Received 20 October 1980; accepted 16 January 1981.

- Piper, P. J. *Pharmac. Ther.* **B3**, 75–98 (1977).
- Kaliner, M., Orange, R. P. & Austen, K. F. *J. exp. Med.* **136**, 556–567 (1972).
- Mezey, R. J., Cohn, M. A., Fernandez, R. J., Januszkiewicz, A. J. & Wanner, A. *Am. Rev. resp. Dis.* **118**, 677–684 (1978).
- Hogg, J. C., Pare, P. D., Boucher, R. C. & Michoud, M.-C. *Can. med. Assoc. J.* **121**, 409–414 (1979).
- Lewis, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3710–3714 (1980).
- Bach, M. K., Brashler, J. R., Johnson, H. G. & McNee, M. L. *Proc. Symp. SRS-A and Leukotrienes* (Research Studies Press, London, in the press).
- Hanna, C. J., Schellenberg, R. R., Pare, P. D. & Bach, M. K. *Clin. Res.* **28**, 702 (1980).
- Bach, M. K. & Brashler, J. R. *J. Immun.* **120**, 998–1005 (1978).
- Bach, M. K., Brashler, J. R., Hammarstrom, S. & Samuelsson, B. *Biochem. biophys. Res. Commun.* **93**, 1121–1126 (1980).
- Bach, M. K., Brashler, J. R., Hammarstrom, S. & Samuelsson, B. *J. Immun.* **125**, 115–117 (1980).
- Van Rossum, J. M. *Archs int. Pharmacodyn.* **143**, 299–330 (1963).
- Collier, H. O. J. & Sweatman, W. J. F. *Nature* **219**, 864–865 (1968).
- Brocklehurst, W. E. *Adv. Drug. Res.* **5**, 109–113 (1970).
- Mathe, A. A. & Strandberg, K. *Acta physiol. scand.* **82**, 460–465 (1971).
- Dahlen, S.-E., Hedqvist, P., Hammarstrom, S. & Samuelsson, B. *Nature* **288**, 484 (1980).



## BOOK REVIEWS

## Genes and culture: an awkward hybrid

Richard Dawkins

THIS little book is an attempt by authors who are sophisticated and accomplished in their own field of evolutionary biology to fit human culture into their scheme of things. I therefore opened it with some anticipation. The first chapter is a sensible and persuasive exposition of the claim, which will be uncontested by all who understand what a modest claim it truly is, that living organisms may usefully be treated as machines programmed to propagate copies of the genes responsible for the programming. The authors note the danger that such language

can lead the unwary imagination to a forest glade where tiny pixies work their will by casting spells on inanimate objects, but such good literary fun should not bother the reader who keeps firmly in mind the true aimless nature of DNA and who remembers that the apparently geneserving behaviour of organisms is an equally aimless consequence of natural selection. Counting on our readers to distinguish the laws of natural selection from the spells of pixies, we will use the image of a purposeful gene to pursue the problem of the origin of learning [p.2].

Ah what touching trust; I cannot resist a wistful sigh. Pixies, hobgoblins, demons and ogres, there will be readers who will find them all lurking in the most unassuming and innocent prose, provided only that "gene" is juxtaposed with language redolent of the boggy world of computers and automation.

Pulliam and Dunford make use of the language of automation in their second chapter, "The Cybernetic Lizard vs. the Toxic Ants". This chapter is part of their long drawn out construction of a niche for learning in the genes-eye view of life. The ethological reader familiar with Konrad Lorenz's classic *Evolution and Modification of Behavior* (Methuen, 1966), or with his later work *Behind the Mirror* (Methuen, 1977), may wonder if the conclusion Pulliam and Dunford reach, important and correct as it is, really needs quite such sustained and protracted rediscovering. Once you start thinking functionally, it certainly makes sense that "the genes, by controlling neural development of the brain, set a learning program that determines how flexible or stereotyped an individual's behavior will be" (p.22). And, moving to the next chapter, "Emotion and Decision", ethologists will need little persuading that evolution of emotions has a lot to do with "genetic selection for decision criteria to

*Programmed to Learn: An Essay on the Evolution of Culture.* By H. Ronald Pulliam and Christopher Dunford. Pp.138. (Columbia University Press: 1980.) \$10, £5.77.

guide learning. Genes exert control over the behaviour of their survival machines by controlling the development of the neural mechanisms that determine which experiences evoke which emotions" (pp. 26-27). Konrad Lorenz and other ethologists should be given some credit for long ago hammering home the idea that "the genes had to define good and bad for their agent", and that "The solution has been to create emotional sensations within the animal agent which makes it want to avoid 'bad' experiences and repeat 'good' ones" (p. 41).

The authors are now ready to move on to their main subject of culture, which they interpret in the common biological sense of traditional behaviour inherited by non-genetic means. Their chapter "Shortcuts to Learning" makes the reasonable point that, among the "values" which a successful gene might be expected to build into its agents' learning programs, is a predilection for copying the ways of respected individuals. "We might sometimes think to ourselves, 'If they say it is good, it is probably good'." An animal can shortcut the normal learning process by picking up useful tricks from other animals. The Japanese macaque stories give their usual good service in this connection and, mercifully, we are spared the now almost obligatory emphasis on the fact that the great Imo was both female and young when she made her brilliant innovations.

After an interlude on economic exchange theory, we move on to "The Mathematics of Culture", a useful summary of the Cavalli-Sforza and Feldman models of cultural inheritance. There are, of course, immense difficulties in the path of any attempt to apply population genetics-like mathematics to cultural inheritance, but the very difficulties themselves are likely to be illuminating and should not be taken as an excuse for funkling the attempt. The possibility that cultural inheritance might, in some sense, be treated as particulate rather than blending is one that intrigues me. This is my way of wondering if "cultural logic conforms to genetic logic" (Chapter 7). In passing I could not help

noting that Pulliam and Dunford faithfully replicate the notorious cultural mutation in the citation of W. D. Hamilton's "The Genetical Evolution of Social Behaviour" (J. Seger and P. Harvey *New Scientist*, 87, 50-51; 1980). This would have been an especially illuminating example for Pulliam and Dunford themselves to have used since it seems not to be a pure, undirected mutation at all. As Seger and Harvey point out, the "mutant" reading, "The Genetical *Theory* of Social Behaviour", has arisen three times independently, and is almost certainly a cultural legacy from that most admired of tribal ancestors, R. A. Fisher, and his "The Genetical Theory of Natural Selection".

The authors may or may not establish their anthropological credentials by going to town on "kinship". An impressive few pages bristle with words such as matrilineal and patrilineal. They coin the interesting phrase, "cultural coefficient of relationship", but actually it turns out to be less interesting than I expected it to be. Instead of being the cultural equivalent of the genetic coefficient of relationship, they apply the term to an awkward hybrid between cultural and genetic notions. They seem unable quite to emancipate themselves from the assumption that genetic "fitness" (capacity to propagate copies of one's genes) has some kind of privileged priority over its cultural equivalent (capacity to propagate one's ideas, etc.). Thus, instead of stressing that, in the domain of cultural inheritance, "brothers" may be defined as those (genetically perhaps unrelated) who have sat at the feet of the same master, they come up with what is really only a kind of culturally weighted version of the genetic coefficient.

If a child's parents have different ideas on a particular issue, let us say that a child learns the father's idea with probability P and learns the mother's idea with probability M. Then, the probability that two siblings learn the same idea from the same parent is  $P^2 + M^2$  . . . This is the probability of cultural identity by descent and can be called the *cultural coefficient of relationship*. Like the genetic coefficient of relationship, the cultural coefficient can be calculated for any pair of relatives.

But if such a cultural coefficient can be calculated for any pair of genetic relatives, can't it also be calculated for any pair of individuals whether genetically related or not?

The last chapter, "The Legend of

Eslok", is rather queer. If I understand it aright, the intention is sensible enough: to use a parable about an imaginary tribe, the Koltinnopinnu to make a point about the conservative role of traditional ritual in preserving ancient wisdom vital for group survival. But the imaginary legend is dragged out with almost Tolkienian attention to gratuitous detail, including learned footnotes on fine points of the fictitious language of the imaginary Koltinnopinno (plurals are formed by

adding "o"). I was left with the feeling that I must be missing the point of the story, since I couldn't see the symbolic (or anagrammatic?) significance of these details. Well, if it comes to that, maybe I missed the point of the whole book: there has to be *some* explanation for my disappointment with it. □

*Richard Dawkins is University Lecturer in Animal Behaviour at Oxford University, and a Fellow of New College.*

## Surprising structures of nucleic acids

C.W. Hilbers

*Nucleic Acid Geometry and Dynamics.* Edited by R.H. Sarma. Pp.424. (Pergamon: 1980.) Hbk £25, \$55; pbk £11, \$24.50.

IN molecular biology nucleic acids are probably the most discussed of all biological macromolecules. Yet, until the early 1970s, information about the structure these molecules can adopt was mainly based upon results from X-ray diffraction experiments on DNA or RNA fibres. During the past decade the situation has rapidly changed. Several events have contributed to this development. Investigators have been able to crystallize the relatively short transfer RNA and determine its structure to atomic resolution. In addition, newly developed organic synthetic methods have permitted the preparation of well-defined fragments of RNA and DNA, so that single-crystal analysis of these oligonucleotides can be performed, and techniques (NMR, T-jump) for the high-resolution study of the structure and dynamics of nucleic acids in solution have become available.

All of these developments have led to surprising results, outstanding among which is the finding that nucleic acids possess a conformational flexibility that can give rise to structures far more complex than the simple double helix.

Some of the excitement pervading the field is reflected in this book. It is a collection of 16 review chapters by different authors in which the structure and dynamics of nucleic acids are discussed in great detail. The first chapter introduces the nomenclature of nucleic acids, and is followed by a discussion of the way in which structural information can be deduced from NMR measurements and a description of the basics of single-crystal X-ray diffraction techniques as applied to nucleic acids. It is somewhat confusing that although different systems of nomenclature are in current use for nucleic acids and are employed throughout the book, only one is presented in the first chapter. Also, one would have preferred the scaling of the NMR spectra to be given in ppm instead of in Hz.

In the following chapters the basic theme of the book unfolds. Sarma describes the spatial configurations of different forms of DNA, and Seeman gives an account of the crystal structures obtained for oligonucleotides. Lerman's proposal in 1961 that aminoacridine dyes bind to DNA by intercalation between stacked bases has now been verified at atomic resolution for a number of planar dyes intercalated in dinucleotides. Accounts of progress in this area are presented by Sobell and by Berman and Neidle. Rich and his co-workers relate the conformational changes in the sugar phosphate backbone of RNA and DNA fragments, introduced by the intercalation of planar dyes, to the changes occurring in the backbone of tRNA resulting from the intercalation of a base between two consecutive bases.

The intercalation of dyes is taken further by Patel, by Krugh *et al.* and by Crothers *et al.* who discuss the information obtained on this subject from solution studies. Fluctuations in nucleic acid structure as monitored by NMR, ESR and tritium and deuterium exchange techniques are covered by Kallenbach and co-workers and by Lerman and co-workers. Finally, theoretical aspects of the structure, flexibility and accessible surface areas are dealt with in the three chapters by Broyde and Hingerty, Olson, and Alden and Kim.

The book is recommended to anyone involved in structural studies of nucleic acids. It contains a wealth of information, written by experts and illustrated with beautifully coloured molecular models many of which were drawn by Irving Geiss. It is not complete, however, in that it does not contain a chapter devoted to the structure and dynamics of tRNA. A potential buyer should also be aware that a large part of the material has already appeared in *Stereodynamics of Molecular Systems*, edited by R.H. Sarma and published by Pergamon. □

*C.W. Hilbers is Professor of Physical Chemistry at the University of Nijmegen, The Netherlands.*

## Stuck on bacteria

N. Sharon

*Bacterial Adherence. Receptors and Recognition*, Series B, Vol.6. Edited by E.H. Beachey. Pp.466. (Chapman and Hall/Methuen Inc.: 1980.) £27.50, \$69.95.

IT is a truism that in order to inhabit or infect a susceptible host, bacteria must be able to "stick" to the tissue surfaces of the host. Surprisingly enough, until the early 1970s little attention was paid to the problem of bacterial adherence to cell surfaces, its molecular mechanism and possible role in host-parasite interactions. The recent explosion of activity in this area may be ascribed, at least in part, to the current intense interest in receptors on cell surfaces, and in their role in intercellular recognition. No wonder that the first treatise on bacterial adherence appears as part of the series on "Receptors and Recognition".

The 14 chapters in the book are all written by experts active in the field, including several of the pioneers who for some years worked almost alone on bacterial adherence. Notable among them are J.P. Duguid from the University of Dundee and R.J. Gibbons from Boston. The first contribution, by I. Ofek and E.H. Beachey, is an excellent overview of the field. In the subsequent 11 chapters, emphasis is either on adhesion to different surfaces (mucosal, dental or of plants, for example), on the adhesive properties of specific organisms (*Enterobacteriaceae*, *Vibrio*, *Neisseria*, for example), or the binding properties of bacterial appendages (fimbriae or pili) and surface polymers (lipoteichoic acid). This loose organization is probably a reflection of the dearth of knowledge of the basic mechanisms involved in adherence. One of the two concluding chapters deals with recognition systems between eukaryotic cells, a topic of relevance to the prokaryote-eukaryote system of bacterial adherence, and the other with prospects for prevention of the association of harmful bacteria with most mucosal surfaces.

From reading the book it is evident that the adherence of many bacteria to epithelial cells is mediated by complementary molecules — adhesins on the bacteria which react with receptors on the cells in a lock and key manner. In the case of *Escherichia coli* and other *Enterobacteriaceae*, the adhesins are lectin-like molecules in the form of fimbriae that can bind reversibly to sugar residues (predominantly D-mannose) on the surface of the epithelial cells, whereas certain mycoplasmas appear to bind in a similar manner to sialic acid residues. Streptococci, on the other hand, appear to bind to cells via their surface lipoteichoic acid. Much remains, however, to be learned about the nature of the adhesins on other bacteria and about the receptors on the tissues, none of which



has been characterized and about which little is known. More importantly, there is a need for conclusive evidence on the role of adherence in natural infection; if the adhesins play a central role in the ability of pathogenic bacteria to launch infection, then the prevention of adhesion should be an effective way to protect against bacterial infections. This could hopefully be achieved by the application of specific receptor analogues or antibodies to the adhesins.

On the whole, the various chapters are well written with not too much overlap. The references are up to date and the tables abundant and clear. There are many illustrations and, with a few exceptions, they are of high quality, including several artistic scanning electron micrographs

which are most impressive. However, the book suffers from the modern malaise of poor indexing — the number of entries is insufficient and even major items such as adhesins, agglutination, mannose, mucus, virulence and yeast cells have not been included.

This is a valuable book, which should be read not only by researchers in the growing field of bacterial adherence, but also by all those interested in infectious diseases and microbial ecology in general, and all students of cell surface receptors and recognition. □

*N. Sharon is the Sadie and Joseph Danciger Professor of Molecular Biology and Head of the Department of Biophysics at The Weizmann Institute of Science, Rehovoth, Israel.*

## Perspective on the flowering of Fermiology

Walter A. Harrison

*Electrons at the Fermi Surface.* Edited by M. Springford. Pp.541. (Cambridge University Press: 1980.) £35, \$85.

THIS *Festschrift* is a fitting tribute to Professor David Shoenberg, whose experimental studies of electrons at the Fermi surface span nearly a half century. In 1937, he pioneered the use of the de Haas-van Alphen effect, which arises from the quantized orbits of electrons moving in a magnetic field, to determine the electronic structure of the semimetal bismuth. When such studies were extended to metals, however, they correlated only poorly with knowledge of other properties of the metal and the contemporary theory of electronic structure. The experimental effect tended to "see" the tiniest pieces of Fermi surface, the abstract surface in momentum space (or wavenumber space) which is the

boundary separating electronic states which are occupied in a solid from those which are empty. Most properties, and available theory, focused upon grosser features, so there was little confrontation of theory and experiment. In the early 1950s it was gradually recognized that the de Haas-van Alphen effect was giving unambiguous geometrical information about the Fermi surface, in contrast to properties such as conductivity which depend upon a muddle of geometry, dynamical properties and scattering. The relevant question became clear: what is the *true* Fermi surface? This, rather than: which model of the electronic structure gives the best fit to experiment?

Lifshitz and Kaganov give historical perspective to this period in the first contribution to the *Festschrift*. Their view was of an arbitrary dispersion law for electrons

which was to be deduced from experiment. Indeed, in 1954 Lifshitz and Pogorelov set up a mathematical scheme to deduce the Fermi surface unambiguously from de Haas-van Alphen data, assuming only that the surface was convex and had a centre of symmetry. In 1956, Gunnarsen, a student of Shoenberg, carried this out in detail for aluminium. He needed to sketch in data for directions of field where the effect disappeared but he could then deduce the presence of cushion-shaped segments of Fermi surface. In hindsight this was misleading, since the real Fermi surface consisted of a set of cylindrical tubes; the missing data were different from what he assumed. A study of lead by Gold, another student of Shoenberg, quickly followed and appeared in 1958. Gold came up with similar data, which he interpreted as somewhat larger cushion-shaped segments — these also turned out ultimately to be a set of tubes. However, Gold had other data and he made a nearly-free-electron construction which led to a very intricate, but qualitatively correct Fermi surface. Though he did not have all of the surfaces matched to the correct data, his study pointed the way to the future: predict a Fermi surface from approximate theory and tune the result by experiment. Lifshitz and Kaganov indicate how difficult it was to accept the nearly-free-electron theory until forced to by its many quantitative successes. One can understand why they were not persuaded by the theoretical arguments in its favour by reading the following contribution on many-body effects by Wilkins; a series of customary assumptions can be stated, but success is their principal justification.

The nature of the study of electrons in metals changed at just the time of Gold's work. Pippard had deduced a Fermi surface for copper, published in 1957, using the anomalous skin effect. He also needed a starting picture, which in his case was the volume of intersection of three mutually perpendicular cylinders. In 1958, Shoenberg succeeded in observing de Haas-van Alphen oscillations in a whisker of copper. By 1960, at the famous Fermi Surface Conference in Cooperstown, New York, the multiply connected Fermi surface of copper had been confirmed and the essential correctness of the nearly-free-electron nature of simple metals established. In addition, other techniques, such as the magneto-acoustic effect, had been developed and applied to Fermiology, though the de Haas-van Alphen effect, because of the high order of the oscillations, remained the champion for precision. This flowering of techniques is reflected in the Lifshitz-Kaganov chapter and, most appropriately, in the simple but complete description of the Shoenberg effect in Pippard's contribution.

The subsequent history of the field is also well documented in this book. Once experiment had confirmed and refined the approximations used in band theory, and

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

David Shoenberg (left) and Brian Pippard discuss a model of the aluminium Fermi surface at the Cooperstown conference, August 1960.

Courtesy of B. W. Roberts, General Electric Research and Development Center

computer technology advanced sufficiently, theory could reliably predict the Fermi surfaces even of transition metals, and ultimately most other properties of these complex systems. The status of this theory is clearly described by Mackintosh and Andersen; indeed, we have the "great calculating machine" postulated by Wigner and Seitz in 1955 which has turned out to be much more interesting than they anticipated. Theory and experiment have moved on to ferromagnetic materials (described in the contribution from Lonzarich), to strained materials (Fawcett, Griessen, Joss, Lee and Perz) and dilute alloys (Coleridge). The de Haas-van Alphen effect has been pushed to extraordinary precision (Coleridge and Templeton) and the waveshapes, as well as the periods, exploited (Higgins and Lowndes). Information about electron scattering times has been extracted (Springford) and can be utilized in transport properties (Chambers).

These are authoritative and well-written articles, with a cohesion arising from the focus on Shoenberg's interests. There is, inevitably, some inconsistency of level and emphasis in such a combined effort and, in this case, an interesting set of repetitions:

the multiple presentations of Fermi liquid theory and of the density functional method reflect the central position which these concepts play in our understanding of metals.

Where will this field go from here? One might have thought that the simple-metal studies had served their purpose by 1960 in leading to pseudopotential perturbation theory which enabled the development of elementary theory of virtually the entire range of metallic properties, and yet that was only the beginning of what Higgins and Lowndes call the "decade of Fermiology". Fermi-surface parameters have now found their way into some physics handbooks so that they become an end in themselves. One could argue that transition metals have been brought into the state of grace of the simple metals and that Fermiology has again served its purpose. However, this field, which is so closely identified with David Shoenberg, seems to carry enough momentum to continue on with no sign of flagging. □

*Walter Harrison is Professor of Applied Physics at Stanford University and was involved in the development of the theory of metallic Fermi surfaces.*

## Polemics for palynologists

Peter D. Moore

*The Northwest European Pollen Flora*, Vol. II. Edited by W. Punt and G.C.S. Clarke. Pp.265. (Elsevier Scientific: 1980.) \$63.50, Dfl.130.

THE second volume in this series will be welcomed by botanists working in such areas as taxonomy, Quaternary palaeoecology, and aerobiology. It follows the first volume in the series in that it consists of a collection of palynological accounts of families which have previously been published in the journal *Reviews of Palaeobotany and Palynology* (Vols 23–28). This practice of reprinting the papers is somewhat inefficient and in future the series will be continued in the form of special issues of the journal dealing specifically with the pollen flora. It is to be hoped that these future volumes will be widely publicized for the sake of those who do not subscribe to the journal and that the special issues will be available separately.

The papers in this volume maintain the high standards set in the first, both in terms of the critical and helpful keys and descriptions, and also the excellent photographs (light and scanning EM). From the palaeoecologist's point of view, particular welcome will be given to the accounts of the Saxifragaceae and Plantaginaceae. In the former, Verbeek-Reuvers takes a fairly conservative position when compared with that of Ferguson and

Webb (*Bot. J. Linn. Soc., Lond.* 63 295; 1970). He emphasizes the problems in distinguishing certain species on the basis of pollen, *Saxifraga aizoides* and *S. oppositifolia* for example. His caution should encourage many palaeopalynologists to go back to the type material and, where necessary, to increase the number of collections of each type to assess the range of variation here described.

In the Plantaginaceae, the differentiation of *Plantago major* and *P. media* is maintained, but the scanning electron micrographs serve only to make one wonder about the criteria upon which it is based. In *P. media*, microechinae (or scabrae) are visible under phase contrast light microscopy, covering the entire exine surface, whereas in *P. major* these are unclear and often totally invisible on the surfaces of verrucae. The SEM pictures of Clarke show that both pollen types possess these scabrae, so it is difficult to understand why they are so obscure on *P. major* under the light microscope.

Other families described in the volume include Solanaceae, Boraginaceae, Valerianaceae, Aceraceae, Haloragaceae and Papaveraceae. It is a book to which all serious palynologists should have ready access. □

*Peter D. Moore is Senior Lecturer in the Department of Plant Sciences, King's College, University of London.*

## Confusing collisions

I. S. Grant

*Deep-Inelastic and Fusion Reactions with Heavy Ions. Lecture Notes in Physics, 117.* Edited by W. von Oertzen. Pp.394. (Springer-Verlag: 1980.) DM49, \$29.

THIS volume contains the proceedings of a symposium held in October 1979 at the Hahn-Meitner-Institut in Berlin. The symposium was chosen as the occasion to dedicate the heavy-ion accelerator VICKSI, which had then been running for almost a year, and to celebrate the centenary of the births of Otto Hahn and Lise Meitner. Hahn and Meitner were associated with the discovery of fission, and it was fitting to devote the symposium to recent studies of the macroscopic behaviour of nuclear matter using heavy ion fusion and deep inelastic reactions.

The experimental papers present evidence for a bewildering variety of different processes which may occur in the collision of two heavy ions. The binary reactions described as fusion fission, fast fission, deep inelastic and quasi-elastic scattering are not clear-cut, but merge into one another in different regions in the space of mass and energy transfer. Nevertheless, it is clear that different phenomena are involved which potentially contain a great deal of information about the dynamical behaviour of nuclear matter. This potential is not easily realized, partly because of the experimental difficulty in recording all the information which may be relevant in each collision, and also because of the obscuring effect of various mechanisms which may contribute to the de-excitation of the composite system or the primary products of the binary reaction. As well as evaporation and fission, there are pre-equilibrium processes which lead to the emission of light particles and which are described in terms (hot spots, jets) which are notions rather than observed entities.

Naturally, the theoretical picture is also confused. One section of the proceedings consists of four theoretical papers all taking a different view of heavy ion collisions. The topic is at the interesting stage when the relationships between its different aspects are not well understood. In a fast-growing field, the proceedings of a symposium held at the end of 1979 are bound to seem a bit dated, but this volume is still a useful survey. □

*I. S. Grant is a Senior Lecturer in the Nuclear Structure Group at Manchester University.*

### Corrigendum

● In the review of *The Principles of Human Biochemical Genetics*, 3rd Edn, by Harry Harris (*Nature* 290, 72) the prices should read: hardback \$65.75, Dfl.135; paperback \$26.75, Dfl.55.